AUG 4 1950
desk

Vol. 41 MAY, 1950 No. 5

PERIODICAL ROOM
GENERAL LIBRARY
OF MICH.

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

An Experimental Projective Technique for the Analysis of Racial Attitudes 257

GRANVILLE B. JOHNSON, JR.

Error Scores as a Measure of Carefulness 279

BENJAMIN FRUCHTER

Choice of Major Subject as Related to American Council Examination Scores and College Grades 292

AUDREY M. SHUEY

The Relation between Intelligence and Achievement Test Results for a Group of Communities 301

ROGER T. LENNON

A Study of High-school Academic Indices as a Criterion for College Admission 309

ISMAEL RODRIGUEZ BOU AND FRANKLIN L. STOVALL

$6.00 per Year • Published Monthly October to May

WARWICK & YORK, INC.

BALTIMORE 2, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

UNIVERSITY OF MICHIGAN LIBRARIES

THE JOURNAL OF Educational Psychology

Established 1910

EDITORIAL BOARD

STEPHEN M. COREY
Teachers College, Columbia University
—*Learning*—

J. B. STROUD
State University of Iowa
—*Individual Differences*—

JACK W. DUNLAP
10 E. 49th St., New York 17, N.Y.
—*Technical, Statistics*—

PERCIVAL M. SYMONDS
Teachers College, Columbia University
—*Mental Hygiene*—

KARL J. HOLZINGER
University of Chicago
—*Factor Analysis*—

MILES A. TINKER
University of Minnesota
—*Reading*—

HAROLD E. JONES
University of California
—*Social Behavior, Child Psychology*—

ALEXANDER G. WESMAN
The Psychological Corporation
—*Tests*—

H. H. REMMERS
Purdue University
—*Attitudes, Teacher Evaluation*—

PAUL A. WITTY
Northwestern University
—*Children's Interests*—

H. E. BUCHHOLZ
Managing Editor
10 E. Centre St., Baltimore 2, Md.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. The JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

Manuscripts may be submitted to any member of the Editorial Board, but the handling of an article will be facilitated if it is sent to that member of the Board who is designated as particularly interested in the phase of psychology dealt with. (Such designations appear after the Editors' names in the list given above.)

Books and other materials for review and correspondence regarding editorial matters should be addressed to The Journal of Educational Psychology, Warwick and York, Inc., Publishers, 10 E. Centre St., Baltimore 2, Md.

Manuscripts should be typed and double-spaced throughout, including quotations, footnotes, and references. In order to attain

(Continued on Inside Back Cover)

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume 41 May, 1950 Number 5

AN EXPERIMENTAL PROJECTIVE TECHNIQUE FOR THE ANALYSIS OF RACIAL ATTITUDES*

GRANVILLE B. JOHNSON, JR.
Arizona State College at Flagstaff

The important rôle which attitudes play in the social development and integration of individuals and societies has precluded much significant research concerning it. The techniques of measurement and evaluation utilized in research on attitudes are as varied as the many problems intrinsic to the study.

A review of the recent literature on attitudes reveals that a large majority of investigations concerning attitudes uses the objective technique in which attitude scales are devised, applied to a specific group, and the results treated statistically. Each attitude is considered a specific characteristic of the subject and it acquires significance through comparison with other members of the selected population. There is most often no attempt to relate the revealed 'attitude' to the total individual. "The resulting conclusions from such a psychologically molecular analysis of attitudes have usually been that people have different attitudes on different subjects, that these attitudes are often internally inconsistent and irrational, that people acquire attitudes and stereotypes as a result of suggestion, propaganda, conditioning and fortuitous experiences which are imposed upon them by outside forces."[1]

Opposed to the molecular analysis of attitudes is the technique utilizing the molar approach. This method of investigation considers an attitude as a manifestation of a total reaction, a

* This article is the first of a series of three concerned with the origin and development of racial attitude. The second study, entitled "The Origin and Development of the Spanish Attitude Toward the Anglo and the Anglo Attitude Toward the Spanish," and the third, "The Relationship Existing Between Bilingualism and Racial Attitude," will appear in later issues of THE JOURNAL.

UNIVERSITY OF MICHIGAN LIBRARIES

part which contributes to the whole and which is determined by the whole. In the study of racial antagonism, for example, it is not sufficient to classify the selected population into pro and con categories. The total personality must be considered as completely as clinical tools will permit in order to determine the bases for the prejudice.

THE PROBLEM

The purpose of the present study was to devise an instrument which could be used in the investigation of racial attitude bases and development by analyzing personality as a whole, thus making it possible to investigate the dynamics of behavior.

MATERIALS AND SUBJECTS

Materials.—(1) The Projective Test of Racial Attitudes composed of six carefully selected pictures relevant to race situations and provocative of projected response was used for the analysis of racial attitudes. There was a hero figure with whom the subject identified himself and thus projected something of the dynamics of his personality into the depicted situations.

The situations conducive to projected response were identical except for the ethnic group variable. For Anglo subjects there were six cards of total Anglo content and six with an Anglo hero in a Spanish situation. For Spanish objects the situations were reversed with a Spanish hero in an Anglo situation in six cards with identical cards without the Anglo element.

(2) A tabulation sheet was used for recording the dynamics of responses.

Subjects.—(1) Anglo: thirty male subjects (four-, eight- and twelve-year levels, N = 90) from a Southwestern public school system.

(2) Spanish: thirty bilingual male subjects (four-, eight- and twelve-year levels, N = 90) from the same Southwestern town.

A further limitation was that members of both groups were required to have lived in this town for the year prior to testing.

THE PROJECTIVE TEST FOR THE STUDY OF RACIAL ATTITUDES

The present investigator through empirical analysis decided that the most successful method of investigating attitude bases and development was by use of the projective technique which

makes it possible to analyze personality on a molar level. This technique distinguishes between opinion and attitude and makes it possible to investigate the dynamics of behavior; these are two factors not found adequately in the traditional objective techniques.

The Selected Pictures of the Projective Test of Racial Attitudes.— When the Projective Test of Racial Attitudes was originally conceived, eighteen pictures were experimentally tested. It was

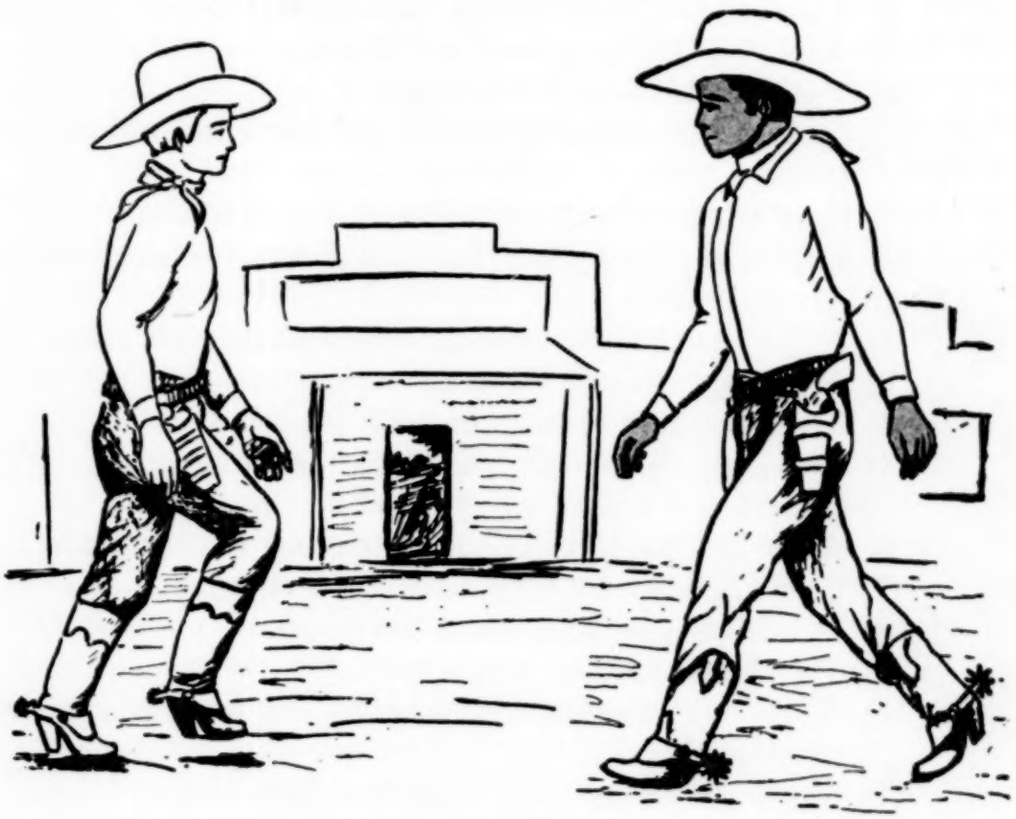

Fig. 1.

found that the six in the present series were adequate in that they covered generally most social situations an individual of the selected ages would ordinarily meet; they fell within the attention span of even the four-year-olds, were of intrinsic interest at each of the selected age levels, and contained life situations with portrayed behavior most easily empathized by the subjects.

The cards are in color and consequently cannot be reproduced here. However, one is reproduced in black and white to show the character of the pictures. (See Fig 1.) They were presented in the following experimentally derived sequence:

UNIVERSITY OF MICHIGAN LIBRARIES

For Spanish Subjects—

1) Spanish and Anglo cowboys approaching each other.
2) Spanish boy running down the street.
3) Spanish boy and Anglo boy holding a bat.
4) Two Spanish boys standing with broken toy between them.
5) Anglo girl crying with Spanish boy observing.
6) Spanish girl possibly inviting Spanish boy to share soda.
7) Two Spanish cowboys approaching each other.
8) Anglo boy running down the street.
9) Two Spanish boys holding bat.
10) Spanish boy standing with Anglo boy with broken toy between them.
11) Spanish girl crying with Spanish boy observing.
12) Anglo girl possibly inviting Spanish boy to share soda.

For Anglo Subjects—

1) Anglo and Spanish cowboys approaching each other.
2) Anglo boy running down the street.
3) Anglo boy and Spanish boy holding a bat.
4) Two Anglo boys with broken toy between them.
5) Spanish girl crying with Anglo boy observing.
6) Anglo girl possibly inviting Anglo boy to share soda.
7) Two Anglo cowboys approaching each other.
8) Spanish boy running down the street.
9) Two Anglo boys holding a bat.
10) Anglo boy standing with Spanish boy with broken toy between them.
11) Anglo girl crying with Anglo boy observing.
12) Spanish girl possibly inviting Anglo boy to share soda.

Administration.—

A) Physical conditions: The physical conditions of the test situation were necessarily controlled. Good lighting was essential because identification by the subject of the characters in the pictures was of utmost importance. It was important, too, to have sufficient table space with no inhibitory stimuli in evidence.

Rapport between the subject and the administrator was necessarily established before testing could proceed. Friendliness and assurance were important in the administrator.

B) Instructions: After rapport had been established, the following information was secured from the subject: (1) age, (2) name, (3) residence in the town (minimum of one year). At this point nothing was mentioned concerning Spanish and Anglo relations.

The directions for administration are as follows: "I am going to show you two pictures (present the Spanish and Anglo stereotypes for mental set) and I want you to tell me which of the two boys you like better—which would you rather be?" Then, "Which do you look like?" After the choice had been made, the following directions were given: "This is not a test. I am just going to show you some pictures and you are going to make up a story about each one—an exciting story—as exciting as you can. There are no right or wrong answers. You have all the time you want. All you are going to do is to make up an exciting story about the pictures I am going to show you."

After the administrator had answered any questions, the first picture was presented.

C) Procedure: 1) Time. (a) The reaction time was recorded. (b) The total time was recorded.

2) Recording. In the upper section of the answer sheet pertinent information was recorded regarding age, name, length of residence in the town with space for further information. The answer sheet was divided into two columns. The responses to the first six pictures were recorded in essence, and verbatim, if particularly significant, in the left column, and the second six in the right column. In this manner, the Spanish-Spanish (or Anglo-Anglo) pictures were directly across from their Spanish-Anglo counterparts. Roman numerals signified the pictures. The reaction time was recorded directly behind the numeral and the total time immediately after the story.

3) Questioning. When necessary inquiry was made concerning the incidents which lead up to the situation depicted, what was happening at the present time, and what would ultimately happen. Questioning was subtle and not leading. It was necessary to repeat that an exciting story was desirable.

After stories had been told for each of the twelve pictures, the subject was asked in direct questioning, how he got along with the members of the other ethnic group; in which were most of his friends; and what he thought of the other ethnic group in

UNIVERSITY OF MICHIGAN LIBRARIES

general. From these data, a correlation between opinion and attitude was made.

4) Length of session. The time required to administer the test generally ranged from thirty to sixty minutes depending on the intelligence, imagination, and stability of the subject.

Principles of Interpretation.—The stories were first read as a whole so as to determine roughly the general trends in the themes.

The method of scoring employed an objective type chart.[2]

I. Effect of the environment on the organism.
   A) Frustrating: rejection, withdrawal.
   B) Neutral.
   C) Helpful: acceptance, friendliness.
II. Reaction of the organism to the environment.
   A) Frustration: insufficiency.
      1. Aggression.
      2. Submission.
   B) Self-sufficiency: emotional stability.
      1. Attainment.
      2. No conflict.
III. Adequacy of the principal character.
   A) Superordination of hero.
   B) Equality.
   C) Subordination of hero.
IV. Ending (hero).
   A) Defeat.
   B) Indeterminate.
   C) Victory, success.
V. Ending (theme).
   A) Unsatisfactory to society.
   B) Indeterminate.
   C) Satisfactory to society.

After the test was administered, it was scored with respect to the four categories, and the chart for scoring, Table I, was marked. Since each picture was identical with its prime number picture except for the race variable, any difference between the 'Spanish-Spanish' and 'Spanish-Anglo' columns were attributed to the single variable (order of administration was controlled). This was done by finding the standard error of the per cent for the total for each year and then finding the critical ratio. The size

Table I.—Tabulation Sheet

Name_____ Age____ Grade_____ Ethnic Group_____ Residence_____ Bilingual Score_____

	Effect of Environ.		React. to Environ.		Adeq. of Pr. Char.		Ending (Hero)		Ending (Theme)	
	Sp-Sp or Sp-Ang		Sp-Sp or Sp-Ang		Sp-Sp or Sp-Ang		Sp-Sp or Sp-Ang		Sp-Sp or Sp-Ang	
	Ang-Ang F N H	F N H	Ang-Ang F S	Sp E Sb	Ang-Ang Sp E Sb	Sp E Sb	Ang-Ang D I V	D I V	Ang-Ang U I S	U I S
Picture 1-1′	A B C	A B C	A B	A B	A B C	A B C	A B C	A B C	A B C	A B C
2-2′	A B C	A B C	A B	A B	A B C	A B C	A B C	A B C	A B C	A B C
3-3′	A B C	A B C	A B	A B	A B C	A B C	A B C	A B C	A B C	A B C
4-4′	A B C	A B C	A B	A B	A B C	A B C	A B C	A B C	A B C	A B C
5-5′	A B C	A B C	A B	A B	A B C	A B C	A B C	A B C	A B C	A B C
6-6′	A B C	A B C	A B	A B	A B C	A B C	A B C	A B C	A B C	A B C
Total										

Reaction Time—Theme

1		7	
2		8	
3		9	
4		10	
5		11	
6		12	
Total			

Total Time—Theme

1		7	
2		8	
3		9	
4		10	
5		11	
6		12	

UNIVERSITY OF MICHIGAN LIBRARIES

of the difference was assumed to be determined by and reflective of the strength of the stimulus to the particular subject. The strength of the stimulus varied from subject to subject since each experience was affect-provoking to the degree that it reflected the individual's total adjustment pattern.

The findings of the Projective Test of Racial Attitudes were then compared with the degree of bilingualism; that is, differences for each of the five categories were compared with the extent of bilingualism.

Scoring the Projective Test of Racial Attitudes.—It is important in the interpretation of protocols derived from projective techniques to ascertain whether or not the subject is identifying himself with one of the characters in his stories since he sometimes does not. In some cases, he manifests an attitude or theory regarding the outside world.

Five principles may be used in determining when the subject is identifying, and with which character in his story:

(a) Other conditions being constant, the subject is likely to identify himself with a character of the same sex, either one of his own age or a previous age.

(b) Other conditions being constant, the subject is likely to identify himself with the central character of the story, that is, the one around whom the story revolves, whose feelings are being expressed, whose behavior is being described, the one who is likely to be described first and figures in the ending.

(c) Other conditions being constant, the subject is likely to identify himself with a character who does not have behavior which is socially unacceptable 'in the subject's eyes.'

(d) The subject is more likely to be identifying with one of the characters when he gives evidence of emotional involvement in the story he tells.

(e) The subject is likely to identify himself with a character whose history is most similar to his own.[3]

In the Projective Test of Racial Attitudes, it was decided that the ethnic group to which the subject belonged determined the hero or principal character. In the Spanish-Spanish and Anglo-Anglo situations, the dominant individual was considered the hero as reflected in the category 'adequacy of the principal character.' (Adequacy of the character with whom the subject identified himself.) Because of the manner in which the scoring

chart was set up, this technique was considered justifiable. Since the present investigation was concerned with prejudice, consistent identification of the subject with one or the other ethnic groups was significant in light of the five principles of identification outlined above.

Following is the frame of reference for scoring the Projective Test of Racial Attitudes. This scoring key was set up before any of the protocols were analyzed. It was necessary, however, to add to it when new problems arose.

KEY FOR SCORING THE PROJECTIVE TEST OF RACIAL ATTITUDES

The hero of the Spanish-Spanish or Anglo-Anglo pictures is the dominant individual. The hero of the Spanish-Anglo situations is the ethnic group to which the subject belongs. If the protocol is divided into a complete Anglo-Anglo and a Spanish-Anglo story, grade both categories. When extra scapegoat (minority group, not including Indians), grade from standpoint of principal character and from standpoint of attitude toward minority group.

Effect of Environment.—Primarily the environment leading up to the present situation.

Helpful: In the Anglo-Anglo or Spanish-Spanish situation, one betters the position of the other. In the Spanish-Anglo situation, the Spanish must help the Anglo or the Anglo must help the Spanish. In the Anglo-Anglo or Spanish-Spanish boy-girl situation, the girl must better the boy's position. If the two characters are in business together—helpful.

Frustrating: Does not tend to set up a problem for the other? Does the Spanish set up a problem for the Anglo in the Spanish-Anglo situation? Hindrance.

Neutral: Neither helpful nor frustrating.

Reaction to the Environment.—Is the effect of the environment frustration provoking or not? Frustration is non-integrative behavior, e.g., fight, flee, or some other overt expression of non-integrative emotion.

Adequacy of Principal Character.—Assume that the ethnic group of the subject is the hero.

Superior: Where the hero's will is imposed on the environment. If in Card II the subject puts the object in a frustrating situation that is superior. In games, the person in most desirable position from subject's standpoint is superior.

Equal: No conflict of will.

Subordinate: Will of other member of the environment is imposed on the hero. Is the hero superior, equal, or subordinate in the environment set up by the other member?

If the boy hits the girl for fun, that is frustration. If he is chastised for act, that is subordination and defeat if the girl is not punished too. Physical, mental, skill, or any other feature is criterion of adequacy.

UNIVERSITY OF MICHIGAN LIBRARIES

Explicit behavior alone is graded except when the subject qualifies behavior. Buying something for girl is not superior.

Defeat in fight equals subordinate. Flees and gets away equals superior. One wins in fight one day and loses next equals equal. Always use one category if possible. If one wins in one thing and loses in another equals superior and subordinate. If one wins in one thing and loses in same thing equals equal.

Ending Hero.—Does the hero end victoriously or defeated with reference to the other member of the picture?

Non-aggression is indeterminant.

If ethnic group I (Hero) has fight with ethnic group II, and wins, that is Victory. Then, if they are both punished for fighting by some member of ethnic group I, that remains Victory. If punished by some member of ethnic group II, that is Defeat.

Ending Theme.—Does the story turn out socially integrated with reference to the mores of society (necessary here to project). Transgression yields retribution. Is the super-ego developed as per the mores of the social group?

Adherence to this frame of reference is essential in the analysis of the protocols.

STATISTICAL ANALYSIS OF THE PROJECTIVE TEST OF RACIAL ATTITUDES

Reliability.—Probably the most valid measure of reliability of scoring technique is to compute the degree of correlation between judgments of two or more trained analysts. In the present investigation, the protocols were scored independently by the investigator and another clinician trained in projective methodology.

TABLE II.—RELIABILITY OF SCORING THE PROJECTIVE TEST OF RACIAL ATTITUDES USING SIXTY SPANISH AND SIXTY ANGLO SUBJECTS DERIVED BY CORRELATING JUDGMENTS OF TWO ANALYSTS

	4 yrs.		8 yrs.		12 yrs.	
	r	or	r	or	r	or
Anglo	.9859	.0031	.9724	.0050	.9691	.0045
Spanish	.9679	.0029	.9466	.0047	.9367	.0111

Table II is a presentation of reliability of scoring technique ascertained by correlating judgments on one hundred and twenty protocols of two trained analysts.

It is noted from Table II that the Spanish protocols yielded slightly less consistency in judgment than the Anglo; and the protocols of the older subjects elicited greater error in scoring than the younger. The former finding was possibly due to the subjectivity (projection) of the two analysts who were both members of the Anglo ethnic group. The latter was probably the result of increased complexity of stories elicited in the older age groups.

A more thorough analysis of scoring consistency was made to discover the reliability of each of the columns in the scoring chart.

The findings indicate no consistent variability. It is the opinion of the present investigator that the derived coefficients of reliability were, to a marked degree, attributable to the effectiveness of the frame of reference and would have approached perfect agreement if the scoring key had been adequately used.

EFFECT OF ETHNIC GROUP OF ADMINISTRATOR ON TEST RESULTS

After the reliability of the Projective Test of Racial Attitudes was established, the effect on test results of the ethnic group to which the administrator belonged was determined. It was realized that the results might be colored by the administrator, particularly if he were a stereotyped member of either of the two groups under consideration.

After two weeks of in-service training including actual practice paralleled by instruction in theoretical background, a Spanish stereotype was allowed to administer a representative sample of tests. A larger sample would have been desirable; but, because of the limited number of available subjects, and the fact that results could possibly be colored by this variable, only one-third of the subjects at the eight-year level in each ethnic group were considered.

An individual with the physiological characteristics of the stereotyped Anglo administered the test to ten subjects in each ethnic group at the eight-year level, and the Spanish administrator to ten others. The difference between the ten Anglo administered

UNIVERSITY OF MICHIGAN LIBRARIES

tests and the ten Spanish administered tests though manifesting inconsistencies (probably because of small sample) demonstrated no constant variation.

As a preliminary experimental procedure, the Anglo investigator administered ten tests which were readministered by the Spanish administrator three months later. The time interval was to remove any effects of practice. The method was necessarily discontinued. The test re-test correlations, though varying in theme, were not altered to a great extent in scored form. The correlations with the small sample were over .80.

With reference to the foregoing analysis, the tentative conclusion was drawn that the ethnic group to which the administrator belonged was not an uncontrolled variable which must be considered in test interpretation.

The Prejudice Score.—It was found that, in order to study the characteristics of the individual protocol, a single score indicating extent of prejudice had to be ascertained. This was done by finding the per cent difference on the tabulation sheet between same group and not-same group categories and adding algebraically. The result was a 'prejudice' score. Analysis of the score, too, brought out the relationships existing between the categories of the scoring chart.

Since the prejudice score was composite of the differences between the per cents of deviate scores of the same group and not-same group categories, and represents a general factor, it was necessary to measure the significance of the contribution made by each difference. By 'deviate scores' is meant those categories which are not neutral, equal, or indeterminate; these latter being non-reflective of positive or negative attitude.

Coefficients of correlation were between each category contributing to the prejudice score minus the weight of the variable category. This procedure was applied at each age level for each ethnic group.

It was noted that, though some categories contributed a negative value at some age levels, the weighted contribution was sufficiently small, (it was demonstrated by experimentation) that little bearing on prejudice score was in evidence.

The prejudice score proved invaluable in the analysis of other variables which demanded individual analysis.

Time.—

(1) Reaction time. The time elapsing between presentation of the stimulus and the subject's first reaction to that stimulus in word association, Rorschach, and other projective techniques, is considered quite important as an indication of emotional arousal and blocking.

One of the features of the Projective Test of Racial Attitudes was to measure the reaction time to the Anglo-Anglo (Spanish-Spanish) pictures and compare it with the reaction time to the Spanish-Anglo pictures. With reference to past use of reaction time, it might be expected that, reflective of increased affect and blocking, the reaction time to the Spanish-Anglo pictures would be significantly greater than reaction time to the like-group situations.

Analysis of Spanish subjects revealed no significant relationship between prejudice and reaction time. Tests of curvilinearity demonstrated no significant deviation from rectilinearity, though there appeared to be a propensity for the level of significance of curvilinearity (x^2Nyx) to increase with age. This would indicate some tendency for Spanish subjects of high and low reaction time to have less prejudice.

In analysis of Anglo reaction time correlated with the prejudice score, it was observed that, though none of the rectilinear correlations are significant, they are uniformly negative. This would indicate that there is a slight though consistent tendency to react more rapidly to emotion provoking stimuli. This is particularly manifest at the twelve-year level.

Analysis of curvilinearity reinforces the tentative conclusions derived with reference to the progressive tendency for those subjects with less prejudice to elicit greater or lesser reaction time. With reference to the relevant literature on the significance of reaction time, no indication of similar conclusions from previous research has been revealed.

(2) Total time. Total time appears particularly at the higher age levels, to be more reflective of emotionality than is reaction time. This conclusion was reached after computing the degree of correlation between prejudice score and total time.

It was noted that as age increases, there was greater relationship existing between the prejudice score and total time.

UNIVERSITY OF MICHIGAN LIBRARIES

Analysis of curvilinear relationships revealed no significant deviation from rectilinearity.

(3) Reaction time and total time. In an attempt to reveal more of the characteristics of reaction time and total time, the relationship existing between the two was determined.

The rectilinear correlation at the four-year level alone is significant ($r = .632$). This would indicate that there is a definite rectilinear relationship between reaction time and total time only at this level.

Curvilinear analysis shows that there is a tendency at the four-year level for those subjects with high and low reaction time to have the least total time. This characteristic, however, becomes progressively less manifest.

Probably the most outstanding feature revealed by the curvilinear analysis is the significant tendency (Eta = .881) at the twelve-year Spanish level for subjects with high and low total time to have low reaction time. This appears to be a recently acquired characteristic of the twelve-year Spanish since there is little indication of this quality at the eight-year level. There is a possibility that this finding may be explained in the following manner—the subject is emotionally aroused and reacts in one of two ways: (1) By responding rapidly to the card, giving a perfunctory, non-revealing response and then going to the next card where he repeats the performance. This would, perhaps, be similar to the evasive nature of the 'cheap whole' to the Rorschach. This hypothesis is given added credence when it is found that the individual subjects are quite consistent in the length of time they require on the individual cards. (2) By immediately seeing the added 'exciting' possibilities in the Spanish-Anglo card and rapidly responding to it. This hypothesis was borne out by subjective evaluation.

The rectilinear correlations appear to be least significant at the four-year Anglo level, ($r = -.04$). This is the antithesis of the Spanish four-year level where the relationship between reaction time and total time was most significant of the age levels, ($r = .63$).

There is also indication, but at a less significant degree than that found in the Spanish subjects, for Anglo subjects with high and low total time to have low reaction time. The possible

explanation for this phenomenon offered for the Spanish subjects may also be applied to the Anglo.

With reference to the findings on reaction time and total time, it may be concluded that reaction time has no significance as an index of prejudice. Total time, however, is more significant, particularly at the higher age levels.

IDENTIFICATION OF SUBJECTS WITH STEREOTYPES

One of the features of administration of the Projective Test of Racial Attitudes was to present the pictures of stereotyped Spanish and Anglo boys and to ask the subject which one he most resembled. This was done to ascertain the rôle of ego in the development of group consciousness and group identification. It was surmised, too, that this procedure might throw light on the hero identification in the stories elicited. As has been pointed out, previous research indicates that ability to identify stereotypes is positively related to the extent of prejudice.

For the purpose of analyzing this feature of the Projective Test of Racial Attitude, the identification of the Spanish and Anglo subjects with the stereotyped Spanish and Anglo drawings was correlated with the individual prejudice score. This was done to ascertain the relationship between identification and attitude.

It is noted that the Spanish and Anglo subjects at the four-year level manifest equal ability to identify themselves with the ethnic group to which they belonged. The Spanish, however, appeared to learn more rapidly than the Anglo, the ethnic group to which they belonged. Only one case in each of the older groups failed to identify himself, though both failures were elicited from subjects who were typically Spanish. An analysis of these two isolated protocols brought out no significant qualities such as prejudice swing which would explain the subject's inability to identify himself correctly.

It was noted that the Anglos, as a result of the age variable, learned to identify themselves less accurately than the Spanish. This is possibly because the physical appearance of the Anglos is less homogeneous than the Spanish. This factor was considered while compiling the data and it was revealed that those older

UNIVERSITY OF MICHIGAN LIBRARIES

Anglos with a high prejudice score who could most conceivably identify themselves with the Spanish stereotype failed to do so.

When the prejudice score was related with the self-identification of the Spanish four-year level, a significant correlation was obtained ($r = -.504$). The four-year level alone could be analyzed because the error of identification at the older age levels was insignificant.

It appears that the higher the prejudice score, the less prone the Spanish four-year-old was to identify himself with his ethnic group. This feature might be explained on the basis (which is borne out by subjective analysis) that the Anglo have a higher economic standing than the Spanish, hence can have more material objects such as better clothes and more toys. This desire appears to be coupled with resentment, as the coefficient of correlation would indicate. Further understanding of this subject may be gained upon analysis of the Anglo four-year level.

It is noted at the Anglo level that there is a progressive tendency for valid self-identification to be related to prejudice until, at the twelve-year level the relationship is marked ($r_{bis} = .43$). Similar to the correlation manifest at the Spanish four-year level, there is a tendency for those four-year Anglos of most accentuated prejudice to identify themselves as members of the Spanish ethnic group, possibly for the same reason suggested for the Spanish behavior. This surmise is supported by a typical statement made by a four-year Anglo whose prejudice score was positive and approximated the mean for members of that age and ethnic group. Upon being asked which of the two stereotypes he most represented, he indicated the Spanish though he was a typical Anglo. He then volunteered the following statement: "I like those boys (indicating Anglo) because they (indicating Spanish) won't play with me. They (Spanish) play a lot better games than the others and are stronger, but they won't play with me because I'm not big and strong enough." The Anglo attribution of superior physical prowess to the Spanish is in evidence throughout the Anglo age levels. Where it is looked upon with admiration by the members of the younger age levels, however, it is considered with contempt by the older. Examples of this shift are as follows:

"I look like (indicating Anglo) and I'd rather be white; if I'd rather be mean and tough, though, I'd rather be (indicating Spanish). But I'd rather be white."

"Most Spanish have better physical ability than whites."

"The Mexicans don't treat you right—they fight all the time. They're tough because they mature earlier than whites, but they're not as smart as whites."

These samples were drawn from the twelve-year Anglo level and the responses were volunteered. An analysis of these specific cases did not demonstrate any marked swing of prejudice: that is, this sensed superiority of the Spanish group was not reflected in greater or lesser bias on the part of the Anglo.

THE STEREOTYPED PICTURE LIKED BETTER

In an attempt to contribute further to the understanding of the prejudice factor, the subjects of both ethnic groups were asked to indicate the stereotype picture they liked better—which one they would rather be like. This choice was then correlated with the prejudice score.

A tendency was found in the Spanish ethnic group for those subjects at the four-year level who were more prejudiced, to like and want to be like the members of the Anglo group. Though this tendency is not significant, it is indicative. There appears to be a reversal in this attitude which is quite manifest by the eight-year level—both the eight- and twelve-year correlations are statistically significant in a position manner. This latter finding would indicate that those older Spanish subjects who preferred the Anglo stereotype and would rather have the appearance of the Anglo possessed a lower than average prejudice factor.

The relationship existing between the choice of stereotype with the prejudice score for the Anglo age levels was made. A comparison of the reaction at the Spanish four-year level with the Anglo four-year level shows a marked difference between groups. Where the prejudice score manifested a negative correlation at the four-year Spanish level, which indicated that those subjects with the greatest prejudice against the Anglo liked and preferred to be like the Anglo, there was a reversal of this finding at the four-year Anglo level which approached statistical significance (critical ratio at .10 level of significance). This would indicate that the Anglos at each age level who preferred the Spanish stereotype were the least prejudiced.

Opinion.—As has been demonstrated in previous research, opinion may not be reflective of attitude. Therefore, it was

UNIVERSITY OF MICHIGAN LIBRARIES

one of the areas of the current investigation to analyze any relationship which existed between opinion and attitude.

An inconclusive tendency was manifest by the Spanish to become more moderate of their opinion of the Anglos with increase in age. A decrease in positive and negative reaction as maturation took place is observable. This trend toward equality, however, does not appear in the Anglo group. Here, dislike of the Spanish increased markedly from the four- to eight-year level with a concommitant decrease in columns headed 'Liked Spanish-Anglo Equally.'

The next step in the analysis of the relationship existing between opinion and attitude was to correlate expressed opinion with the prejudice score of the Projective Test of Racial Attitudes.

At the Spanish four-year level, there is a negative relationship between opinion and attitude, ($r = -.20$), but by the eight-year level little relationship exists, ($r = .096$). Evidently, the significant positive correlation at the oldest age level, ($r = .71$), is the product of recent development.

The same trend was manifested when the Anglo age levels were analyzed. The tendency, however, was not as accentuated as that found in the Spanish group.

Analysis of older groups would be important to ascertain the relationship between opinion and attitude at more sophisticated levels.

INTERNAL CONSISTENCY OF THE PROJECTIVE TEST OF RACIAL ATTITUDES

Each card in the Projective Test of Racial Attitudes was selected on the basis of its ability to cover a particular area of the subject's experience range. The environmental variability was quite wide, and varied in complexity.

The contribution which each picture makes to the end result might be valuable in determining its effectiveness, but a picture that has a low correlation with the group might be invaluable because it is loaded with an important but esoteric factor. In the Thematic Apperception Test, as has been previously indicated, it has been found that the responses to the pictures cannot be inter-correlated because the pictures are uneven in the type and amount of response they elicit and because the method includes isolating factors from context.

This rationalization is, no doubt, true, but the present investigator was interested in (1) noting the existence of a general factor and (2) ascertaining the consistency of behavior from situation to situation with reference to racial attitudes. The technique employed intercorrelates the response to each picture with the responses to the other five for each of the ten categories for each ethnic group at each age level. The aim, then, of correlating each picture with the others was to note the existence of a general factor and to determine the consistency of subject reaction.

When the study had been accomplished and the results analyzed it was concluded that each picture of the Projective Test of Racial Attitudes had a positive relationship with the other pictures of the test. The degree of relationship, though differing, displays no consistent variation resulting from the influence of age and ethnic group. Though there is a marked general factor, each picture makes a specific contribution to the total prejudice score.

SUMMARY

Probably the most successful method of investigating attitude bases and development is by utilization of the projective technique which makes it possible to analyze personality on a molar level.

Six pictures were experimentally selected and proven to be adequate. They constitute the Projective Test of Racial Attitudes.

Techniques of administration and interpretation were standardized and rigidly controlled to eliminate uncontrolled variables which would invalidate results.

Reliability of scoring the Projective Test of Racial Attitudes was found to be high at each age level for both ethnic groups.

The effect on test results of the ethnic group to which the administrator belonged was analyzed. Upon analysis of this problem, the tentative conclusion was drawn that the ethnic group to which the administrator belonged had little bearing on test results and was not an uncontrolled variable which must be considered in test interpretation.

It was found that, in order to study the characteristics of the individual protocol, a single score indicating extent of prejudice had to be ascertained. This was accomplished by finding the

UNIVERSITY OF MICHIGAN LIBRARIES

per cent difference of the tabulation sheet between same group and not-same group categories and adding algebraically. The result was a 'prejudice' score, which proved invaluable in the analysis of other variables which demanded individual analysis.

One of the features of the Projective Test of Racial Attitudes was to measure the reaction time to the Anglo-Anglo (Spanish-Spanish) pictures and compare it with the reaction time to the Spanish-Anglo pictures. No significant rectilinear or curvilinear relationship between racial attitude and reaction time was noted, but there was a tendency in both ethnic groups for those subjects with less prejudice to elicit greater or lesser reaction time. No indication of similar conclusions from previous research has been revealed.

Total time appeared, particularly at the higher age levels, to be more reflective of emotionality than was reaction time. This conclusion was reached after correlating prejudice score with total time.

The relationship between reaction time and total time was determined and it was found that there was a tendency for those subjects with high and low total time to have low reaction time.

One of the features of administration of the Projective Test of Racial Attitudes was to present the pictures of stereotyped Spanish and Anglo boys and to ask the subject which one he most resembled. It was noted that the Anglos, as a result of the age variable, learned to identify themselves less accurately than the Spanish. This is possibly because the Anglo physical appearance is less homogeneous than the Spanish. It was revealed, however, that those older Anglos with a high prejudice score who could most conceivably identify themselves with the Spanish stereotype failed to do so. There was indication that the higher the prejudice score, the less prone the four-year-olds of both ethnic groups were to identify themselves with their own ethnic group. A progressive tendency was manifest in both ethnic groups for valid self-identification to be related to prejudice score until, at the twelve-year level the relationship was marked.

A tendency was found for those Spanish subjects at the four-year level who were most prejudiced, to like and want to be like the members of the Anglo group. Though this tendency was not significant, it was indicative. There appeared to be a

reversal in this attitude which was quite manifest by the eight-year level—both the eight- and twelve-year correlations were statistically significant in a positive manner. This latter finding would indicate that those older Spanish subjects who preferred the Anglo stereotype and would rather have the appearance of the Anglo possessed a lower than average prejudice factor.

A comparison of the reaction at the Spanish four-year level with the Anglo four-year level showed a marked difference. Where the prejudice score manifested a negative correlation at the four-year Spanish level, which indicated that those subjects with the greatest prejudice against the Anglo liked and preferred to be like the Anglo, there was a reversal of this finding at the four-year Anglo level which approached statistical significance (critical ratio at .10 level of significance). This would indicate that the Anglos, at each age level, who preferred the Spanish stereotype were least prejudiced.

Since opinion may not be reflective of attitude, one of the areas of the current investigation was to analyze any relationship existing between opinion and attitude. An inconclusive tendency was manifested by the Spanish to become more moderate of their opinion of the Anglos with increase in age. A decrease in positive and negative reaction as maturation took place was observable. This trend toward equality, however, did not appear in the Anglo group. Here, dislike of the Spanish increased markedly from the four-year to the eight-year level with a concomitant decrease in columns headed 'Liked Spanish Better' and 'Liked Spanish-Anglo Equally.'

The next step in the analysis of the relationship existing between opinion and attitude was to correlate expressed opinion with the prejudice score of the Projective Test of Racial Attitudes. It was found that, at the four-year level in the Spanish group, there was a marked negative relationship between the two factors under consideration, but by the eight-year level little relationship existed. Evidently, the significant positive correlation at the oldest age level was the product of recent development.

The same trend was manifested when the Anglo age levels were analyzed. The tendency, however, was not as accentuated as that found in the Spanish group.

Each card in the Projective Test of Racial Attitudes was

UNIVERSITY OF MICHIGAN LIBRARIES

selected on the basis of its ability to cover a particular area of the subject's experience range. The environmental variability was quite wide, and varied in complexity.

The present investigation was concerned with, (1) noting the existence of a general factor, and (2) ascertaining the consistency of behavior from situation to situation with reference to racial attitudes. The technique employed intercorrelates the response to the other five for each of the ten categories for each ethnic group at each age level. After this procedure was followed, it was concluded, that each picture of the Projective Test of Racial Attitudes had a positive relationship with the other pictures of the test. The degree of relationship, though differing, displayed no consistent variation resulting from the influence of age and ethnic group. Though there was a marked general factor, each picture made a specific contribution to the total prejudice score.

REFERENCES

1) David Krech. "Attitudes and Learning: A Methodological Note," *Psychological Review*, Vol. 53, No. 5, September, 1946, p. 291.

2) The idea of scoring the thematic responses in this manner was procured from Ruth Clark (Genetic Psychology Monographs, 1944, Vol. 30, p. 47). H. A. Murray, in a recent letter to the present investigator, made some specific suggestions, and the present investigator, after advice from L. I. O'Kelly, and through experimentation, developed the scoring system in its present form.

3) Julian R. Rotter. "Thematic Apperception Tests: Suggestions for Administration and Interpretation," *Journal of Personality*, Vol. 15, September, 1946, p. 83.

ERROR SCORES AS A MEASURE OF CAREFULNESS

BENJAMIN FRUCHTER*
University of Texas

I. INTRODUCTION

There are a number of occupations that require extremely careful and meticulous work. An example is the navigator of the modern long-range bomber. He may be called upon to plot a course to hit a tiny dot of land in the broad Pacific. Miscalculation may result in the loss of expensive equipment and the lives of himself and the other crew members.

In order to test the hypothesis that there is a trait of carefulness and that a factor of carefulness can be isolated and identified, four tests were constructed. These are paper-and-pencil tests similar to some of the tasks of the navigator such as careful measuring and plotting on maps and charts and reading values from complicated scales.

These tests were administered to three hundred fifty-four unclassified aviation students in January, 1945. The directions did not stress either speed or accuracy to the exclusion of the other. Two scores were obtained from the same administration of each test—the number of correct responses and the number of errors. It was observed that the error scores had considerable range and variability. It was also observed that when the correct and error scores of a test were correlated the coefficients were relatively low, indicating considerable independence of the two scores. This suggested the possibility that the error scores might be functionally different from the correct-response scores and should be treated separately in an analysis. In order to examine the hypothesis that these tests measured a common trait of carefulness it was determined to submit them to factor analysis along with a number of reference tests.

II. THE TESTS†

1–2. Directional Plotting.—This is a test of the speed and accuracy with which an examinee can locate points and estimate

* The writer wishes to express his appreciation to Roger W. Birkman and William L. Grafton for assistance with the factor analyses reported in this study.

† For a more complete description of the tests see J. P. Guilford (Ed.),

UNIVERSITY OF MICHIGAN LIBRARIES

directions on a chart. The coördinates of two points on a 30 × 30 chart are given. The task is to (1) locate the positions of the two points without making any marks on the chart, (2) determine the direction of the second position from the first in terms of a twelve-point compass rose shown at the side of the chart. The rights score is variable 1 and the wrongs score is variable 2.

3–4. Complex Scale Reading.—This test consists of a chart containing seven vertical scales. For each item the examinee is given values on two of the seven scales. By placing a straight edge through these points, the examinee is to read the corresponding value on a given one of the remaining five scales. The rights score is variable 3 and the wrongs score is variable 4.

5–6. Plotting.—This test consists of a 25 × 25 chart and a compass rose. The examinee is given a starting point in terms of coördinates and told to move so many squares west, so many squares southeast, so many squares east, etc., and then to read the final position in terms of coördinates. The rights score is variable 5 and the wrongs score is variable 6.

7–8. Plotting Accuracy.—This test is a variation of the Plotting test. The same kind of chart is used, but three orientation compasses are shown as compared with one in the plotting test. The points on these are marked with letters rather than directions. The examinee is instructed which compass rose to use in each item. The rights score is variable 7 and the wrongs score is variable 8.

In order to make up a matrix for analysis eleven tests, selected from the classification test battery routinely administered to all aviation students in the AAF, were intercorrelated with the eight scores (four rights scores and four wrongs scores) obtained from the carefulness tests. The list includes several apparatus classification tests, because they showed high correlations with the carefulness tests. A brief description of each of these reference tests follows:

9. Rotary Pursuit.—This is a modification of the Koerth Pursuit Test. The examinee tries with his right hand to keep a prod in contact with a metallic spot on a phonograph-type disc which is rotating at the rate of one revolution per second. While

Printed Classification Tests, Washington: United States Government. Printing Office, 1947, pp. 680–686.

doing this, there is a second simultaneous task for the left hand which requires the examinee to keep one of two keys closed in correspondence with one of two lights.

10. Two-hand Coördination.—This test consists of a rotating black disc on which is a small brass 'target' disc. The target disc follows an irregular path. The task of the examinee is to keep a microswitch in contact with the target disc as long as possible. The movement of the microswitch is controlled by turning two handles similar to those on a lathe. The handles turn in the opposite direction from those of a conventional lathe, so that those having had previous lathe experience will not have an undue advantage.

11. Complex Coördination.—This is a serial, choice-reaction-time test in which each stimulus is one of thirteen spatial patterns of three lights each. In correspondence with each stimulus pattern, the correct response is a unique adjustment of imitation stick-and-rudder controls.

12. Rudder Control.—In this test the examinee sits in a mock cockpit of an airplane. His own weight throws the seat off balance unless he applies correction by means of a rudder-control mechanism. The score is the total time he keeps the cockpit pointed directly at a target light straight ahead.

13. Discrimination Reaction Time.—This test was designed as a test of speed of decision and reaction. There are four stimulus patterns, each consisting of a pair of lights, one red and one green. Corresponding to each stimulus pattern is a microswitch, the four switches being arranged in a diamond-shaped pattern. The position of each switch, upper, lower, right, and left, is associated with a corresponding direction of the red light with respect to the green light. The task of the examinee is to push the switch associated with a given lighted pattern. The score is the time it takes the examinee to push the correct switch.

14. Finger Dexterity.—This test consists of forty-eight square pegs in square holes. Each peg can be grasped by means of a thick circular button at its top. The examinee lifts each peg from its hole, turns it 180° clockwise and resets it in its hole. The score is the total number of pegs turned in the time allowed.

15. Speed of Identification.—This is a paper-and-pencil test in which the silhouette of a plane must be matched with the

UNIVERSITY OF MICHIGAN LIBRARIES

identical silhouette in a rotated position. The stimulus silhouettes are presented in groups of four and the silhouettes with which they are to be matched in groups of five, one being a mislead. The test is printed directly on an IBM answer sheet.

16. Numerical Operations I.—This is a simple arithmetic test printed directly on an IBM answer sheet. It consists of simple problems of addition and multiplication with answers provided. The examinee's task is to mark the given answer as either correct or incorrect.

17. Mechanical Principles.—This test is similar to the well-known Bennett and Fry Mechanical Comprehension Test. The items are pictorial representations of simple mechanical principles. The answer choices are stated verbally.

18. Reading Comprehension.—This test is composed of eight paragraphs concerning which thirty-six questions are asked. The items were designed to test ability to make valid inferences from the reading material as well as ability to answer more direct questions about content.

19. Arithmetic Reasoning.—This test consists of thirty arithmetic-reasoning problems which require a minimum of formal mathematical training to solve.

III. THE ANALYSIS

The rights scores and wrongs scores were obtained from the same administration and were separated for analysis. Failure to do so would result in experimentally dependent variables due to the fact that both scores had been taken from the same set of responses.*

Two matrices were prepared for analysis. One contains the four rights scores of the carefulness tests plus the formula scores (R − .25W) of the eleven reference tests. The other contains the four wrongs scores plus the eleven reference tests. These reference tests had been analyzed several times previously and their factorial content for this population is relatively well established. They were included to facilitate cross-validation

* For an analysis of these tests in a single matrix see J. P. Guilford (Ed.), *op. cit.*, pp. 686-694. In that solution four doublet factors, one for each carefulness test, were obtained. They are spurious and resulted from the fact that both rights, and wrongs scores for any given test were derived from the same sample.

of factors in the two carefulness batteries as well as with factors isolated in previous studies with these reference tests.*

The wrongs scores were reflected for correlational purposes. Intercorrelations of the carefulness tests and correlations of the carefulness tests with the reference tests are based on a sample of three hundred fifty-four unclassified aviation students. The intercorrelations of the reference tests were not available for this particular sample but are based on other comparable samples of unclassified aviation students with a total N of 1,920. The intercorrelations of all the variables involved in the two analyses are shown in Table 1.

Each matrix of intercorrelations was factored by the centroid method. Six factors each were extracted from the rights and wrongs matrices. These factors were rotated graphically to "psychologically meaningful positions" in accordance with the principles of simple structure and positive manifold. Inspection of the factor plots indicated little evidence of obliqueness and an orthogonal reference frame was maintained. Table 2 gives the rotated and centroid loadings for the rights battery and Table 3 the corresponding values for the wrongs battery.

IV. INTERPRETATION OF FACTORS

Rotated Factor I is interpreted in both batteries as a numerical-facility factor.

Test	Loading	
	Rights	Wrongs
Numerical Operations I	.78	.61
Complex Scale Reading	.57	.09
Plotting	.61	.05
Plotting Accuracy	.61	.18
Arithmetic Reasoning	.44	.51
Directional Plotting	.39	−.07

All the rights scores of the carefulness tests have numerical content. This is to be expected from the nature of the tasks.

* Ibid., chapter 28.

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 1.—CORRELATION MATRIX*

Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Directional Plotting (R)																			
2. Directional Plotting (W)	48																		
3. Complex Scale Reading (R)	57	05																	
4. Complex Scale Reading (W)	10	28	43																
5. Plotting (R)	46	12	55	24															
6. Plotting (W)	07	20	13	39	42														
7. Plotting Accuracy (R)	46	09	51	20	71	19													
8. Plotting Accuracy (W)	04	14	14	34	26	41	46												
9. Rotary Pursuit	36	−10	37	−05	49	−12	48	−06											
10. Two-Hand Coordination	44	−04	44	01	51	−03	44	−04	33										
11. Complex Coordination	42	19	30	11	35	08	32	09	38	48									
12. Rudder Control	20	20	04	07	07	−01	05	−06	32	30	32								
13. Discrimination Reaction Time	46	14	43	12	32	08	35	13	18	23	36	07							
14. Finger Dexterity	16	03	17	00	24	05	22	05	33	27	35	11	26						
15. Speed of Identification	27	10	28	03	34	05	29	08	12	14	21	05	21	16					
16. Numerical Operations I	38	04	49	07	48	05	52	14	05	−04	05	−03	16	14	11				
17. Mechanical Principles	34	36	18	23	11	10	12	14	21	35	29	27	20	04	13	−16			
18. Reading Comprehension	37	14	34	09	19	08	23	11	03	05	16	−01	21	04	11	15	31		
19. Arithmetic Reasoning	45	17	46	16	32	08	34	16	−02	06	15	−01	23	01	04	35	27	50	

* Decimal points omitted. Wrongs scores are reflected.

TABLE 2.—RIGHTS SCORE BATTERY
(Decimal points omitted throughout)

Test	Rotated Loadings						Centroid Loadings						
	I	II	III	IV	V	VI	I	II	III	IV	V	VI	h^2
1. Directional Plotting (R)	39	28	44	38	09	20	75	−10	19	09	−12	06	63
3. Complex Scale Reading (R)	57	24	28	16	11	42	73	−27	−09	−06	−16	20	68
5. Plotting (R)	61	30	06	07	54	19	75	−14	−41	−16	06	−16	80
7. Plotting Accuracy (R)	61	24	10	08	50	16	73	−20	−34	−14	06	−17	74
9. Rotary Pursuit	13	60	02	07	17	29	52	33	−25	−08	−16	18	51
10. Two-Hand Coordination	05	50	18	10	49	05	58	37	−11	−17	13	−12	55
11. Complex Coordination	00	43	18	44	31	03	58	36	06	18	11	−07	52
12. Rudder Control	−08	52	12	17	−04	−12	26	44	10	08	−24	−11	35
13. Discrimination Reaction Time	14	13	22	44	17	27	52	−05	10	22	16	16	39
14. Finger Dexterity	11	37	−09	29	08	11	36	21	−18	22	10	05	27
15. Speed of Identification	11	00	03	22	26	25	34	−06	−09	13	24	12	22
16. Numerical Operations I	78	−04	−04	19	−03	−03	40	−54	−23	19	−28	−22	62
17. Mechanical Principles	−17	23	56	09	16	08	39	24	42	−22	09	07	44
18. Reading Comprehension	16	−09	58	11	06	17	40	−26	41	−15	07	10	43
19. Arithmetic Reasoning	44	−09	62	06	02	−02	47	−41	39	−18	−10	−14	61

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 3.—WRONGS SCORE BATTERY
(Decimal points omitted throughout)

Test	Rotated Loadings						Centroid Loadings						
	I	II	III	IV	V	VI	I	II	III	IV	V	VI	h^2
2. Directional Plotting (W)	−07	−10	14	16	41	36	37	−18	−25	13	17	−29	36
4. Complex Scale Reading (W)	09	−01	05	00	14	60	38	−36	−32	−14	07	01	40
6. Plotting (W)	05	−05	−02	07	−07	64	31	−38	−31	−30	−09	−03	43
8. Plotting Accuracy (W)	18	−04	06	10	−15	57	34	−39	−15	−30	−15	04	41
9. Rotary Pursuit	12	56	−04	23	13	−14	33	50	13	−10	14	15	43
10. Two-hand Coordination	−09	65	28	13	−02	00	43	49	−08	06	−21	24	54
11. Complex Coordination	06	57	19	34	13	16	63	37	−03	−08	−04	07	55
12. Rudder Control	−04	42	00	−03	42	06	32	30	−22	12	31	13	37
13. Discrimination Reaction Time	18	19	21	41	00	13	48	08	19	−10	−12	−10	31
14. Finger Dexterity	18	32	−09	44	01	02	36	28	20	−30	05	−06	34
15. Speed of Identification	04	10	07	34	01	07	28	10	10	−11	−07	−17	15
16. Numerical Operations I	61	−01	−01	07	00	02	22	−24	43	−16	18	20	39
17. Mechanical Principles	−18	22	53	11	37	24	55	08	−27	42	−09	−10	58
18. Reading Comprehension	26	−11	57	18	08	05	41	−23	28	33	−19	−07	45
19. Arithmetic Reasoning	51	−06	56	−02	12	11	46	−37	33	34	−05	19	61

The wrongs scores do not have significant loadings on the numerical factor.

Rotated Factor II is interpreted in both analyses as a psychomotor-coördination factor.

Test	Loading	
	Rights	Wrongs
Rotary Pursuit	.60	.56
Rudder Control	.52	.42
Two-hand Coördination	.50	.65
Complex Coördination	.43	.57

Four of the six apparatus tests have significant loadings on this factor. A similar factor has been identified several times previously in analyses of the aircrew classification battery.* Several of the carefulness tests rights scores have loadings approaching significance on this factor, reflecting the high correlations of these tests with some of the apparatus tests.

Rotated Factor III is interpreted as a general-intellectual factor, possibly a combination of the general-reasoning and verbal factors usually obtained separately.

Test	Loading	
	Rights	Wrongs
Arithmetic Reasoning	.62	.56
Reading Comprehension	.58	.57
Mechanical Principles	.56	.53
Directional Plotting	.44	.14

Only one of the rights scores of the carefulness tests has an appreciable loading on this factor.

Rotated Factor IV is identified as a psychomotor-precision factor in both analyses.

* Ibid., p. 816.

UNIVERSITY OF MICHIGAN LIBRARIES

Test	Loading	
	Rights	Wrongs
Discrimination Reaction Time	.44	.41
Complex Coördination	.44	.34
Directional Plotting	.38	.16
Finger Dexterity	.29	.44

A similar factor has been identified in previous analyses of these apparatus tests.* The Directional Plotting, rights score, apparently involves some finger dexterity.

Rotated Factor V, rights battery, is interpreted as a spatial-reference factor involving judgments of change of distance and direction. Rotated Factor V, wrongs battery, is identified as a visualization factor.

Test	Loading	
	Rights	Wrongs
Plotting	.54	−.07
Plotting Accuracy	.50	−.15
Two-hand Coördination	.49	−.02
Complex Coördination	.31	.13
Rudder Control	−.04	.42
Directional Plotting	.09	.41
Mechanical Principles	.16	.37

There is, so far as the writer knows, no precedent for the factor interpreted as spatial-reference and the identification is accordingly tentative. The loading of Directional Plotting, wrongs score, on the visualization factor is of interest. Presumably the ability to visualize the compass rose on the chart helps to keep down the error score on that test.

Rotated Factor VI, wrongs battery is interpreted as a care-

* Ibid., p. 817.

fulness factor. The corresponding factor in the rights battery is not interpreted since only one test has a large loading on it.

Test	Loading	
	Rights	Wrongs
Plotting	.19	.64
Complex Scale Reading	.42	.60
Plotting Accuracy	.16	.57
Directional Plotting	.20	.36

The wrongs scores of the four carefulness tests have significant loadings on the factor from the wrongs battery. Since a trait of carefulness was postulated in the construction of these tests the factor is so identified. The sixth factor from the rights battery has but a single carefulness test loaded on it. This, together with the fact that the second and third highest loadings on it are psychomotor tests (see Table 2), makes it unlikely that it is a carefulness factor too. The extent to which the carefulness factor is common to error scores in other tests remains to be determined.

V. DISCUSSION

The major question which this study has attempted to answer is whether the error score of a speeded test contains sufficiently unique variance to justify its treatment as a separate variable. The error scores of the four experimental tests included in this analysis did define a new factor. Had the error scores not been analyzed separately, the general conclusion would have been that no new factor resembling carefulness could be found in these carefulness tests.

Thurstone* has isolated a similar factor in his analysis of perceptual tests. The tests were scored either per cent wrong or per cent omitted. He says of this: "The factor involved may be concerned with accuracy or caution. More data should be available on a variety of test material for this factor before it

*L. L. Thurstone, "The Perceptual Factor," *Psychometrika*, 3, 1938, pp. 1–17.

UNIVERSITY OF MICHIGAN LIBRARIES

can be identified. This finding does suggest, however, that the relative frequency of errors may represent a unique trait."

The previous analysis of these data, referred to above, in which both the rights and wrongs were included in the same analysis, confirms the results that the error scores do yield a unique factor even when the rights scores from the same tests are included in the matrix.* Rights scores of the Plotting and Plotting Accuracy tests did show slight, but probably not significant loadings (.22 and .19, respectively) with the carefulness factor.

It may be assumed safely that, if correlations of rights and wrongs scores are not high, a fuller picture of the true functions measured by a test can be obtained by analyzing the scores separately than by analyzing formula scores. The results also imply that many an error has possibly been committed by combining rights and wrongs scores in the same formula. Unless the two are factorially similar the result may be very different than had been intended by the test-maker. In this connection, it is in order to suggest that if rights and wrongs scores from a test are considered in factor analysis or in composite predictions, they should be derived from separate forms in order to avoid spurious correlations.

Another result of interest is the high correlation of the carefulness tests with some of the apparatus tests. The carefulness tests had higher correlations with the apparatus tests than with the other paper-and-pencil tests with which they were correlated. This fact gives considerable encouragement to proponents of the belief that many factors appearing in apparatus tasks can be duplicated in printed tests.

SUMMARY

In order to test the hypothesis that a trait of carefulness could be identified, four tests similar to some of the tasks performed by the aerial navigator were constructed. These are paper-and-pencil tests of plotting on maps and charts and reading values from complicated scales. The tests were administered to three hundred fifty-four unclassified aviation students in the AAF and separate correct-response and error scores were obtained for each test.

It was observed that the error scores had considerable range

* Guilford, *op. cit.*, p. 691.

and variability and low correlations with the correct-response scores. It was decided to intercorrelate and analyze the scores to determine whether a carefulness factor could be identified.

Eleven paper-and-pencil and apparatus tests which had been previously taken by the aviation students in the course of their classification were added as reference variables. Two matrices were prepared for analysis, one consisting of the correct-response scores of the carefulness tests plus the eleven reference variables and the other of the error scores plus the reference variables.

The two matrices were factored and the factors rotated to simple structure. Numerical (I), psycho-motor coördination (II), general-intellectual (III), and psycho-motor precision (IV), factors were isolated in both analyses. The rights battery yielded a spatial-reference factor for the fifth factor (R-V). The fifth factor on the wrongs battery (W-V) appears to be a visualization factor. The sixth factor on the rights battery had an appreciable loading on only one test and was not interpreted. The sixth factor on the wrongs battery (W-VI) is the carefulness factor whose isolation was sought.

The major conclusions are:

1) The error scores of the four experimental tests define a new factor.

2) Had the error scores not been analyzed separately no new factor resembling carefulness could have been found in these tests.

3) If correlations of rights and wrongs scores are not high, a fuller picture of the functions measured by a test can be obtained by analyzing the scores separately than by analyzing scores arrived at by scoring formulas such as R − .25W etc.

4) If rights and wrongs scores from a test are considered in composite predictions or factor analyses, they should be derived from separate forms in order to avoid spurious correlations.

UNIVERSITY OF MICHIGAN LIBRARIES

CHOICE OF MAJOR SUBJECT AS RELATED TO AMERICAN COUNCIL EXAMINATION SCORE AND COLLEGE GRADES

AUDREY M. SHUEY

Assisted by Barbara Edgar
Randolph-Macon Woman's College

INTRODUCTION

While a number of investigators have reported a relation between intelligence and occupation, very few have been concerned with intelligence as related to the field of specialization at the college level. One psychologist, in attempting to determine whether or not intelligence increases from the freshman to senior year, presented some findings pertinent to this problem. After testing a Lawrence College freshman class on the 1934 edition of the American Council on Education Psychological Examination, Flory[1] grouped the students according to their choice of major subject and subsequently combined those in related fields. He found the Fine Arts group to rank first on the A. C. E., followed in order by the Language Arts group (English, French, German, Spanish, Latin, and Greek), the Science group (Biology, Chemistry, Geology, Mathematics, and Physics), and the Social Science group (Economics, Government, History, Psychology, Philosophy, Religion, and Sociology). Hartson[2] observed that Oberlin College freshmen who subsequently majored in the Mathematics—Science division scored higher on the average in numerical computations when tested by the Ohio State University Psychological Examination, while those freshmen who were to major in the Language-Literature groups showed a greater initial ability in handling verbal relationships.

The original purpose of the research to be reported in this paper was to determine whether or not women in a Liberal Arts College who major in one field differ in general intelligence from those who select another area for concentration. Finding that some departments did differ significantly from others in the general intelligence of their majors, we thereupon made a survey of grades by departments, since it seemed possible that students

of somewhat lesser ability would have some difficulty in making their grades during their first two years and would select for further study those divisions of instruction where the grading was somewhat lenient. If this be the case one would not expect to find the various departments of different colleges ranking in the same way with respect to the general intelligence of their majors.

PROCEDURE

A score on the A. C. E. was used as the criterion of general intelligence in this study. For a number of years, the particular college studied has required all new students to take this test at a specific time during their first month of college residence. From the files of the Psychology Department were selected the names and test scores of all freshmen tested in the nine years between 1939 and 1947. The students tested in 1939 had been given the 1939 edition; those tested in 1940, the 1940 edition; . . . and those in 1947, the 1947 edition. Over the period of years they were all tested in the same large lecture rooms, at the same time of day, and by members of the Psychology Department. As far as the writers know, no new student who was on the campus during the testing period has ever been absent from the test.

Of the 2012 freshmen tested during the nine-year period, 1279 were in the college during their second year, selected a major field of concentration during April of this year, and indicated that they would return for the junior year. Of this number, a total of twenty-seven students selected German, Greek, Philosophy, Physics, or Religion for specialization, and were excluded from our survey because of the small numbers. The total gross scores of the remaining 1252 freshmen were all transmuted into the equivalent scores of the 1943 edition and categorized according to choice of one of thirteen major fields.

In determining average grades earned by students in the thirteen divisions of instruction under consideration, we included all students whose surnames began with A to L, inclusive, and who were in the college during their freshman and sophomore years between 1939 and 1947. For all divisions of instruction except Music and Art we have included only grades for six-hour courses, since the great majority of courses—other than in

UNIVERSITY OF MICHIGAN LIBRARIES

these two departments—in the first two years are year-courses allowing six semester hours of credit.* A preliminary study of 1340 consecutive grades of students entering college between 1939 and 1942 indicated that only thirty grades had to be omitted by this procedure. These thirty totalled sixty-three hours of credit as contrasted to the 7860 hours included in the 1310 grades. By our method of selection therefore we have probably included all but one per cent or less of the credits earned in the eleven divisions of instruction.† In Art and Music, all grades, regardless of hours' credit, were included with the exception of Choral Music which we excluded from our tabulations.

Ordinarily a student was given one final grade for the year's work in a course (six hours); however, if the two semesters' grades were recorded separately, the second grade, if not the same as the first, was arbitrarily taken as the grade. Furthermore, if a student dropped a given year course at the end of the first semester or at any other time with a record of D or F, the D or F was included just as if she had remained in the course throughout the year.

RESULTS

In Table 1 are included the number of majors in each of the thirteen divisions of instruction from 1939 to 1947, the mean gross scores and SD's of these students tested on the A. C. E. at the beginning of the freshman year, and the percentile ranks of these means as derived from the national norms. The A. C. E. means range from 133.50 for those students who later selected French as a major, to 119.07 for those who subsequently majored in Biology, the respective percentile ranks of the groups ranging from 86 to 68. The Music majors were most variable in their A. C. E. scores, as indicated by a standard deviation of 22.83, followed by English, History, and French; the SD's of the other

---

* In the Art Department the courses are primarily four-hour courses. In Music likewise the great majority of classes open to freshmen and sophomores are four-hour courses.

† The omitted grades were from the following divisions of instruction: English, French, History, Latin, Mathematics, Political Science, and Spanish.

nine groups were very similar, all of them being between 17 and 18.

No significant differences were found among the freshman scores of groups who subsequently elected to major in French, Mathematics, Chemistry, Spanish, Art, Psychology, and Latin,

TABLE 1.—MEAN GROSS SCORES, STANDARD DEVIATIONS, AND PERCENTILE RANKS OF FRESHMEN TESTED BY THE A. C. E. FROM 1939 TO 1947 AND GROUPED ACCORDING TO SUBSEQUENT CHOICE OF MAJOR[1]

Major	N	Mean	SD	Percentile Rank[2]
Art	31	127.81	17.16	80
Biology	101	119.07	17.32	68
Chemistry	119	129.98	18.04	82
Economics & Sociology	151	120.93	17.74	71
English[3]	299	121.65	19.30	72
French	54	133.50	19.07	86
History	100	121.68	19.20	72
Latin	58	126.33	17.52	77
Mathematics	104	131.02	17.86	83
Music	33	126.00	22.83	77
Political Science	57	119.37	17.86	69
Psychology	84	127.68	17.82	79
Spanish	61	128.41	17.28	80

[1] All individual scores were transmuted into equivalent scores for 1943. All choices of major subjects were made in the Spring of the students' sophomore year; however students who changed their major during their junior or senior year were placed according to their final choice.

[2] Norms based on 31,980 freshmen from two hundred forty-one colleges.[7]

[3] The departments of English Literature and English are combined in this study. Sociology is combined with Economics since for the greater part of the years covered they were included within a single department.

at the one, two or five per cent level of confidence. Likewise none of the differences between the means of the groups who in their second year chose to major in History, English, Economics and Sociology, Political Science, and Biology proved to be significant at any of the three levels of confidence. However,

UNIVERSITY OF MICHIGAN LIBRARIES

the French, Mathematics, Chemistry, Spanish, and Psychology majors proved to have made significantly higher means (one per cent level) than the students who selected Biology, Political Science, Economics and Sociology, and English. The means of the French, Mathematics, and Chemistry majors were also significantly higher than the History mean (one per cent level), while the Spanish and Psychology means were significantly higher than the History mean at the five per cent level of confidence. The means of the Art and Latin majors were significantly above those of Biology, Political Science, and Economics and Sociology at either the two or five per cent level of confidence. Music majors were found to be neither significantly higher nor lower than any other group of majors at the three levels of confidence employed.

In Table 2 are included 7997 grades of freshman and sophomore students, earned between 1939 and 1947 and separated according to the thirteen divisions of instruction under consideration. As will be observed from the table, the mean grade ranges from 3.72 in Art* to 3.18 in French, Psychology, and Spanish; the standard deviations ranging from 1.17 in French to 0.78 in Biology and English.

The divisions of instruction may be conveniently divided into two groups: the eight 'higher-grading' and the five 'lower-grading.' The higher-grading divisions of Art, Latin, Music, Political Science, English, Economics and Sociology, Biology, and History were all significantly above (one per cent level of confidence) the five divisions of French, Spanish, Psychology, Mathematics, and Chemistry. None of the lower-grading divisions differed significantly from one another, but among the higher-grading groups the means of Art and Latin were significantly above (one per cent level) those of History, Biology, Economics and Sociology, and English. The Art mean grade was also higher than that earned in Political Science (one per

---

* Probably one would expect the Art and Music grades to average higher than those in the other divisions of instruction, since these two are the only 'purely elective' courses among the thirteen divisions. In order to graduate a student must have successfully completed courses in English; History; a foreign language or languages; Chemistry, Mathematics, or Physics; Biology or Psychology; and Economics, Sociology, or Political Science.

cent level) and higher than the Music mean (five per cent level of confidence).

It will be observed that the five groups whose means were the lowest on the A. C. E. all selected their major fields of concentration from among the eight higher-grading divisions of instruction. Likewise five of the seven groups who scored relatively

TABLE 2.—MEANS AND STANDARD DEVIATIONS OF FRESHMAN AND SOPHOMORE GRADES FROM 1939 TO 1947 GROUPED ACCORDING TO DIVISION OF INSTRUCTION[1]

Division of Instruction	N	Mean[2]	SD
Art	160	3.72	0.92
Biology	589	3.41	0.78
Chemistry	569	3.25	0.95
Economics & Sociology	450	3.42	0.79
English	1389	3.45	0.78
French	750	3.18	1.17
History	1007	3.40	0.93
Latin	503	3.61	0.90
Mathematics	760	3.23	1.04
Music	379	3.52	0.99
Political Science	328	3.47	0.82
Psychology	341	3.18	0.91
Spanish	772	3.18	1.04

[1] Including only grades of students whose surnames began with A to L and who were in college for two or more years.

[2] Letter grades of A, B, C, D, and F were arbitrarily numbered 5, 4, 3, 2, and 1, respectively.

high on the A. C. E. subsequently selected their major field from among the five lower-grading divisions of instruction. Only in the case of the Art and Latin departments do we find relatively high-scoring A. C. E. groups choosing their major fields from the higher-grading divisions. The coefficient of correlation by the Product-moment method between mean A. C. E. scores of majors and the mean grades of freshmen and sophomores in the same division of instruction is −.44.

SUMMARY AND CONCLUSIONS

1) All freshmen in a southern Liberal Arts college for women were given the American Council Examination from 1939 to 1947. Among these students tested there were 1279 who were enrolled in the college during their sophomore year, selected a major field for concentration, and indicated their intention of returning for their junior year.

2) Twelve hundred fifty-two of the sophomores selected one of thirteen divisions of instruction for their major. The freshman A. C. E. scores of this number were transmuted into equivalent scores of the 1943 edition.

3) The seven higher-scoring groups, determined by their means on the A. C. E., were: French, Mathematics, Chemistry, Spanish, Art, Psychology, and Latin. None of these groups differed significantly from one another.

4) The five lower-scoring groups, determined by their A. C. E. means, were: History, English, Economics and Sociology, Political Science, and Biology. No significant differences were obtained among these means.

5) The Music majors, the most variable of the groups, did not differ significantly from any of the other twelve groups.

6) Five of the higher-scoring groups, *i.e.*, French, Mathematics, Chemistry, Psychology, and Spanish majors, earned significantly higher test means than the five lower-scoring groups.

7) Approximately 8000 grades secured from 1939 to 1947, the great majority of which indicated six semester hours of work, were tabulated according to the thirteen divisions of instruction. These were selected by recording the freshman and sophomore grades of all students whose surnames began with A to L inclusive and who attended the college for two or more years.

8) The grades of eight divisions of instruction were significantly above those of the other five, on the average. The five lower-grading divisions were those of French, Mathematics, Chemistry, Psychology, and Spanish.

9) Except for the departments of Art and Music, in which none of the courses are required and to which special talents might be expected to direct students, and Latin, there appears some evidence of an inverse relation between mean A. C. E.

score of majors and mean grade earned by beginning students in the field. The divisions of instruction in which freshman and sophomore students were graded lowest on the average, *i.e.*, Chemistry, Mathematics, Spanish, Psychology, and French, were those that had majors who ranked on the average between the 79th and 86th percentiles on the A. C. E.; whereas the departments that graded their freshmen and sophomores somewhat higher, *i.e.*, Biology, Economics and Sociology, English, History, and Political Science, had majors who on the average ranked between the 68th and 72nd percentiles on the A. C. E. It would appear therefore that choice of major subject bears some relationship to students' A. C. E. scores and to the average grades of those enrolled in the more elementary courses.

10) By combining 'related fields' as Flory did, we find our results to be in agreement with this investigator only insofar as the Social Science group scored lowest of the four divisions. Since the ranking of the other three divisions is at variance with that of Flory, it is quite probable that the agreement in the Social Science division is fortuitous.

REFERENCES

1) C. D. Flory. "The intellectual growth of college students." *J. Educ. Res.*, 1940, 33, 443–451.

2) L. D. Hartson. "Does college training influence test intelligence?" *J. Educ. Psychol.*, 1936, 27, 481–491.

3) A. M. Shuey. "Improvement in scores on the American Council Psychological Examination from freshman to senior year." *J. Educ. Psychol.*, 1948, 39, 417–426.

4) L. L. Thurstone and T. G. Thurstone. "Psychological Examinations, 1940 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

5) ——— ———. "Psychological Examinations, 1941 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

6) ——— ———. "Psychological Examination for College Freshmen 1942 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

7) ——— ———. "Psychological Examination for College Freshmen 1943 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

8) ——— ———. "Psychological Examination for College Freshmen 1944 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

9) ——— ———. "Psychological Examination for College Freshmen 1945 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

10) ——— ———. "Psychological Examination for College Freshmen 1946 Norms." *Amer. Counc. Educ. Stud.*, Washington, D. C.

THE RELATION BETWEEN INTELLIGENCE AND ACHIEVEMENT TEST RESULTS FOR A GROUP OF COMMUNITIES[1]

ROGER T. LENNON

Director, Division of Test Research and Service,
World Book Company

What is the relation between the intelligence level of the elementary-school population of a community and its achievement as measured by standard achievement tests? Is this relation the same from grade to grade or does it tend to increase or decrease with additional schooling? Does this relationship vary from subject to subject and, if so, in what manner? These are questions of considerable theoretical importance and not without practical significance, particularly with reference to such matters as the provision of separate achievement test norms for communities or schools of varying intelligence levels. This paper presents data on the relationship between achievement and intelligence as measured by standard tests, for successive grades and for some important subjects in the elementary school.

SOURCE OF DATA

The data reported were obtained in connection with the standardization of the new (1946) edition of the *Metropolitan Achievement Tests.* They are based on the results for seventy of the two hundred fifty communities which participated in the standardization program; these seventy are communities in which all public school pupils in grades II through VIII (or IX, depending on the organization of the system) were tested, and exclusive of very small communities whose results are based on such few cases as to be of very limited dependability. The communities, however, are for the most part, small ones; about two-thirds are of less than ten thousand population. They are drawn from thirty-six states. The pupils in the various grades took the appropriate battery of both the *Pintner General Ability Test* series, and of the *Metropolitan* series, as listed in Table 1.

[1] Paper read at the April, 1949, meeting of the Eastern Psychological Association, Springfield, Massachusetts.

All tests were administered within two months of the opening of schools in the fall of 1946. The figures reported are based on an analysis of a random sample of approximately twenty-five per cent of the total number of cases in each community; for this random sample the mean IQ for each grade in each community was computed, together with the mean standard score in each subtest of the *Metropolitan Achievement Tests* for each grade. It is the relationships among these community averages which are here discussed.

TABLE 1.—BATTERIES OF *Metropolitan Achievement Tests* AND *Pintner General Ability Test* USED IN VARIOUS GRADE LEVELS IN METROPOLITAN STANDARDIZATION PROGRAM

Grade	*Metropolitan Achievement Tests*	*Pintner General Ability Test*
II	Primary I: Word Picture, Word Meaning, Word Recognition, Numbers	Pintner-Cunningham Primary Test, Form C
III	Primary II: Reading, Word Meaning, Arith. Fund., Arith. Prob., Language Usage, Spelling	Pintner-Durost Elementary Test: Scale I (Picture Content), Form B
IV	Elementary: Reading, Vocabulary, Arith. Fund., Arith. Prob., Language Usage, Spelling	Pintner-Durost Elementary Test: Scale II (Reading Content), Form B
V–VI	Intermediate: Reading, Vocabulary, Arith. Fund., Arith. Prob., English, Literature, History, Geography, Science, and Spelling	Pintner General Ability Test: Intermediate Form B, Verbal Series
VII–IX	Advanced: Reading, Vocabulary, Arith. Fund., Arith. Prob., English, Literature, History, Geography, Science, Spelling	Pintner General Ability Test: Intermediate Form B, Verbal Series

The basic data of this paper are summarized in Tables 2 and 3. Table 2 presents the correlations between the average intelligence level (mean IQ) and average achievement level for each grade and subject. In all grades except grade IX these correlations are based on 70 communities; in grade IX they are based on 39 communities. Table 3 presents for each grade the means and standard deviations of the achievement and intelligence test distributions; it is to be noted that these are for distributions of community means, not of individual scores.

SUMMARY OF FINDINGS

Examination of the data presented in Table 2 indicates that all relationships between intelligence and achievement test scores reported are positive; at grade IV and above, with a few exceptions, they are significantly so. (For N = 70, as here, r must equal about .30 for significance at the 1% level.) This is entirely in accord with expectation, although the magnitude of the correlations at the lower grades is perhaps not so great as would have been anticipated. The correlations range from .11 between IQ and what we have called Arithmetic Fundamentals (really Numbers) at the second grade level, to .86 between IQ and Vocabulary score at the eighth grade level; the median value for all correlations is approximately .60. That is to say, in general there is a marked tendency for communities which are above average in IQ to be above average also in achievement, as measured by these tests.

If we turn our attention to the manner in which these relations vary from grade to grade, we note a marked tendency for the correlations to increase as one goes up through the grades; this is clearly revealed by inspection of the right-hand column of Table 2, in which the average value of the correlations for the respective grades is shown. These data indicate that at the second- and third-grade levels, knowledge of the average intelligence test scores for pupils in a community would enable one to improve the prediction of the community's achievement status but slightly, if at all. At the upper grades—VII, VIII, and IX—on the other hand, the tendency is much more pronounced for the communities with high average IQ's to be also the high-achieving communities.

TABLE 2.—CORRELATIONS BETWEEN COMMUNITY MEAN IQ (*Pintner General Ability Test*) AND MEAN STANDARD SCORE ON SUBTESTS OF *Metropolitan Achievement Tests*

Grade	Subject										
	Reading	Word Meaning	Arith. Fund.	Arith. Prob.	Spelling	English	Literature	History-Civics	Geography	Science	Aver.,* all subj.
II	.34[1]	.33	.11[2]								.29
III	.42	.42	.19	.30	.27						.33
IV	.61	.75	.19	.41	.26						.48
V	.68	.75	.25	.44	.37	.58					.54
VI	.72	.81	.13	.43	.62	.60	.63	.50	.67	.62	.60
VII	.79	.85	.42	.50	.54	.66	.74	.63	.70	.67	.67
VIII	.85	.86	.65	.65	.61	.52	.77	.71	.71	.68	.72
IX	.84	.82	.57	.55	.56	.61	.67	.69	.72	.53	.67
Aver.,* all Grades	.66	.74	.36	.48	.49	.60	.70	.64	.70	.63	

[1] Value for both Word Picture and Word Recognition Tests.
[2] Numbers Test.
* Averaged via *z* transformation.

TABLE 3.—MEANS AND STANDARD DEVIATIONS OF COMMUNITY MEAN STANDARD SCORES, BY GRADE, ON SUBTESTS OF THE *Metropolitan Achievement Tests*

Grade	Subject																			
	Reading		Word Meaning		Arith. Fund.		Arith. Prob.		Spelling		English		Literature		History-Civics		Geography		Science		IQ	
	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ
II	112.5	5.8[1]	121.2	6.2	137.2	3.2[2]															100.7	7.2
III	141.2	5.7	142.6	5.9	137.8	2.5	149.4	2.6	142.7	6.7											100.7	5.5
IV	164.9	7.1	165.7	6.6	154.7	5.1	165.8	4.1	163.7	7.7	164.3	7.1									98.3	6.9
V	180.7	8.1	179.2	7.6	171.6	6.3	178.6	4.9	179.6	6.2	177.9	6.5									100.7	6.0
VI	194.7	7.4	196.0	8.0	195.4	10.4	198.2	6.8	196.6	6.3	195.8	8.2	197.3	7.1	199.1	7.4	197.9	7.6	199.5	8.7	99.5	6.0
VII	204.3	7.5	204.7	7.3	220.2	9.4	211.2	7.2	210.1	6.9	205.5	7.5	206.4	8.3	207.8	7.0	208.5	9.1	208.7	8.5	98.9	5.9
VIII	214.9	7.1	216.0	7.2	242.7	12.3	223.8	9.3	228.6	8.4	222.4	10.0	216.9	8.6	218.0	8.4	216.9	9.9	224.5	10.7	99.5	5.9
IX	221.2	6.4	224.2	6.8	257.4	14.5	233.3	9.4	239.0	9.3	233.2	10.2	222.7	6.8	224.8	7.4	223.2	9.2	234.0	9.1	100	5.5

[1] Word Recognition Test.

[2] Numbers Test

This apparent change in the relationship between upper and lower grades may be to some extent an artifact, reflecting differences in the content of the intelligence tests used. In the second grade, the *Pintner-Cunningham Test* was used, and in the third grade, Scale I, the Picture Content Scale, of the *Pintner-Durost Elementary Test* was used. Both of these are non-reading tests and measure abilities that bear little resemblance to school work; whereas the tests used in grades IV to IX are verbal tests, whose subtests include, among others, arithmetic problems, vocabulary, opposites, and number sequence tests, which are obviously influenced by school experience. Hence, it is not surprising that the intelligence and achievement test results for grades IV to IX should agree more closely than in the earlier grades. However, this would seem to be but part of the story, for even between the fourth and the ninth grades, in which range the same intelligence test was used, there is a steady increase in the degree of correspondence between the two sets of scores. One could argue that the intelligence test results are becoming progressively more dependent on, more a reflection of the influences of, schooling; but there is nothing in the data to show in which direction causative influences, if any, are operating.

A second hypothesis would be that the increasing correlation between average IQ and average achievement with increasing schooling is in part a function of greater variability in either measure as one goes up through the grades. The data in Table 3 lend some support to this belief. The variability of community mean IQ's remains quite constant throughout the grades, to be sure, but there is a clear tendency for the communities to become more variable in achievement at the higher grades, as examination of the standard deviations reveals.

What of the relation of community mean IQ to community performance in the separate subjects? For communities, as for individuals, there are marked differences among subjects at all grade levels. In general, the greatest correlation tends to be between the reading measures and the intelligence measures; at the seventh-, eighth- and ninth-grade levels, these correlations almost approach the magnitude that one would expect to find between two measures of intelligence or between two reading tests. Thus, knowing the average intelligence level of pupils in these grades, as measured by the tests used, one could estimate

with considerable accuracy their average reading achievement, or vice versa. The correlations tend to be smallest between IQ and what we may term the 'skill' subjects, namely, arithmetic and spelling; for Arithmetic Fundamentals in grades II to VI, knowledge of intelligence is of little help in predicting achievement. At the lower grades the correlations between Arithmetic Fundamentals and IQ are significantly smaller than the correlations between Arithmetic Problem Solving and intelligence, but this is not the case in grades VII to IX. Between the reading tests, on the one hand, and the 'skill' tests on the other, are the so-called 'content' subjects—Literature, History and Civics, Geography and Science; the correlations between community means in these subjects and intelligence range from .50 to .77.

The attempt to relate achievement measures to intelligence test results has been of persistent concern to test-makers and test-users—and legitimately so, for the desire to evaluate an individual's performance in terms of his 'capacity' is a praiseworthy one. Accomplishment quotient techniques and cognate devices represent one approach to this problem, albeit scarcely a satisfactory one. The provision of separate norms for the interpretation of scores for individual pupils of varying IQ levels, as is done for certain achievement tests, represents another approach. Some writers have suggested that it would be desirable to provide achievement test norms for school systems of varying intelligence levels on the ground that such norms would permit the school administrator to compare achievement in his system with achievement in systems whose pupils are similar in intelligence. This suggestion assumes that variation in the average intelligence level is associated with variation in achievement to a degree sufficient to make it worth while for practical purposes to establish differentiated norms; as mentioned in the opening sentences of this article, the data herein reported are relevant to this question.

The data presented in Table 2 indicate, in the writer's judgment, that in grades II through V, at least, the relationships between the intelligence and achievement levels of a community, with the single exception of those for reading, are not sufficiently large to warrant the establishment of differentiated norms for school systems of varying average intelligence levels. In grades VI through IX the correlations, with the exception of those for

arithmetic, spelling and English, are of such magnitude that one may expect communities scoring high in intelligence tests also to do well on achievement tests with considerable consistency. If the magnitude of this relationship were the only factor to be considered in determining the desirability of differentiated norms, one could argue in favor of such norms at the upper grades except for the subjects named.

There are, however, other elements to be considered in deciding, both on practical and theoretical grounds, whether separate norms of this character are desirable. From a practical standpoint, the provision of multiple sets of norms is expensive, both in terms of the cost of obtaining adequate data and in terms of making the additional normative data available; and it is doubtful that there is sufficient sophistication on the part of the ordinary test-user to make it likely that such differentiated norms would be generally understood and properly used. On theoretical grounds, we might cite the problem of the validity of most currently available intelligence tests as measures of 'capacity' in any real sense, and the corollary problem of the community of function measured by present intelligence and achievement tests. Even if the practical problems in the construction and use of differentiated norms could be coped with successfully, the theoretical justification for such norms seems, to the writer, dubious to say the least.

A STUDY OF HIGH-SCHOOL ACADEMIC INDICES AS A CRITERION FOR COLLEGE ADMISSION

ISMAEL RODRÍGUEZ BOU and FRANKLIN L. STOVALL

Superior Educational Council
University of Puerto Rico
Rio Piedras, Puerto Rico

INTRODUCTION

One of the many problems with which the University of Puerto Rico is faced is that of selecting wisely and well its student body from among the numerous individuals who desire to enroll. This problem exists even now and, without a doubt, will become more serious in the future. At present, the University and other institutions of higher learning on the Island are not able to accommodate all persons who desire to matriculate. Some kind of a screening process seems necessary in order that these institutions shall be able to admit those people who can profit to the greatest extent from college training.

In such a screening process many factors must be taken into consideration in order that those individuals who must be screened out or eliminated (if such be the case) will be the individuals who are least likely to succeed at the college level. One of the factors which has carried most weight in the selection of students by institutions of higher learning has been and still is the factor of high-school grades or previous academic record. The University of Puerto Rico prior to the school year 1946–47 used this factor as a sole criterion for selection.

How reliable is such a criterion? To what degree of accuracy is it possible to predict college grades from high-school grades? Are letter grades of 'A,' 'B,' 'C,' etc. from the various high schools of Puerto Rico equivalent? Do those individuals with the same letter grade averages in high school possess approximately the same degree or amount of preparation for college work?

In order to obtain more accurate and reliable answers to these and similar questions it was decided to make a comprehensive and scientific study of the problem. Accordingly, a study was sponsored by the Superior Educational Council of the University of Puerto Rico with Dr. Ismael Rodríguez Bou, Permanent

Secretary, as Director. This paper is in the form of a summary report* of the study—the purpose, methods and technique, the presentation and analysis of data, the conclusions and, finally, the recommendations.

PURPOSE

The purpose of this study was two-fold:

1) To evaluate the high-school academic record as a criterion for admission to college. With what degree of accuracy is it possible to predict grades in the University of Puerto Rico from grades made in high school?

2) To determine the relative importance of letter grade averages as given to students by the teachers of the various high schools of the Island. Is an 'A' or 'B' grade in one high school equivalent to an 'A' or 'B' grade in some other high school? Are students from different high schools with the same letter grade averages equally prepared for academic work in college?

METHODS AND TECHNIQUES

In the initial stages of this study, the grades made in the University by eight hundred fifty-three freshmen just out of high school (fifty-seven different schools) who were enrolled in the University during the year 1945–46 were obtained along with their respective grades in high school. The University grades made by each student in all courses completed during the 1945–46 school year were averaged in order to get an average grade for each of the eight hundred fifty-three freshmen included in the study. Similarly, an average of high-school grades was computed for each student. In computing the University and high-school averages an 'A' was given the value of 4, a 'B' the value of 3, a 'C' the value of 2, a 'D' the value of 1, and an 'F' a value of 0. The lowest high-school grade point average of any student was 2, for students with less than a 'C' average in high school are not admitted to the University (except veterans who pass a special examination).

These averages were then compared and a coefficient of correlation computed. A regression equation was worked out by means

* A complete comprehensive detailed report of this study has been published as English and Spanish Editions by the Superior Educational Council of Puerto Rico. Copies are available upon request.

of which University grade point averages were predicted from high-school grade point averages. The predicted grades were then compared directly with the grades actually made in the University and the extreme variations noted. An attempt was made to relate these extreme cases to particular high schools.

Some months later the study was extended to include an additional group of one thousand one hundred fifty-seven students who had matriculated in the University for the school year 1944–45 and who had come directly from the high schools of the Island. The same methods and techniques employed in an analysis of the grades of the eight hundred fifty-three students were used in dealing with this last group of one thousand one hundred fifty-seven students.

Also, it was possible still later to study in the same manner two hundred forty-eight and two hundred thirty-four students who entered in 1944–45 and 1945–46, respectively, the College of Agriculture and Mechanical Arts of the University of Puerto Rico, located at Mayagüez. Then one hundred eight students who were freshmen at Polytechnic Institute in 1945–46 were included. Finally, the combined total of two thousand six hundred students was studied.*

DATA FROM 2,600 CASES

Averages of high-school academic indices and University grade points for the several class groups (a total of 2,600 cases) are given in Table I along with the standard deviations and coefficients of correlation between high-school indices and grades made in college. Since high-school graduates who have less than a 'C' average in all their high-school work are not admitted to the University of Puerto Rico and to Polytechnic Institute, the high-school academic indices for the above groups range from 2.00 to 4. In effect, then, the students admitted to college in Puerto Rico come from the upper or 'right' half of at least approximately a normal distribution representing the total of all high-school graduates for the years involved. Grades made in

* The data here deal only with the combined total of all cases and represent graduates of seventy-three different high schools in the Island, both public and private. The total, however, includes six students at Polytechnic Institute who graduated from high school outside of Puerto Rico.

college, however, range from 0 to 4 and follow a normal curve of distribution.

Some interesting comparisons are possible by an examination of Table I. Columns 2 and 3 give the means or arithmetic averages of high-school and college indices. The mean high-school

TABLE I.—MEANS, SIGMAS, AND COEFFICIENTS OF CORRELATION COMPUTED FOR THE SEVERAL GROUPS

Groups	Means		Sigmas		
	High School	University Grades	High School	University Grades	r's
University (853)	2.74	2.20	0.52	0.67	0.57
University (1157)	2.69	1.97	0.53	0.78	0.51
Total: 2010	2.71	2.07	0.53	0.74	0.54
College of Mayagüez (248)	2.48	2.18	0.54	0.60	0.52
College of Mayagüez (234)	2.45	2.01	0.59	0.79	0.70
Total: 482	2.47	2.10	0.57	0.70	0.62
Polytechnic Institute (108)*	2.93	2.21	0.51	0.50	0.54
Grand Total: 2600	2.67	2.08	0.53	0.73	0.57

* Includes six students who graduated from high schools outside of Puerto Rico.

index for students enrolling in the University of Puerto Rico (2,010 cases) is 2.71, for the College of Agriculture at Mayagüez, 2.47 (482 cases), and for Polytechnic Institute at San Germán, 2.93 (only 108 cases for one year). The average high-school index for all 2,600 cases (including six students at the Institute

who graduated from high school outside of Puerto Rico) is 2.67, or a low 'B' average. The mean college index for the University is 2.07; for the College of Agriculture, 2.10; for the Institute, 2.21; and for all combined, 2.08. In columns 4 and 5 of Table I are given the corresponding standard deviations.

In the last column of Table I the coefficients of correlation between high-school and college indices are listed. The averages of these coefficients were computed by squaring each coefficient, securing an average of these squares, and extracting the square root. The correlation for the University and the Institute are the same, .54. The correlation for the College of Agriculture is the highest, .62. The coefficient of correlation for all 2,600 cases combined into one large sample is .57. All of these coefficients are positive and satistically significant, but they are quite low for purposes of prediction. When regression equations were worked out, the error of prediction was so large that for practical purposes it is difficult to predict accurately college grades from a knowledge of the high-school grades. This means then that high-school indices as a criterion for the selection of college students is not very reliable, especially in regard to students with a low 'C,' 'C,' or high 'C' average.

Table II lists the seventy-three public and private high schools which constituted the source for the student groups studied. The total number of students and the average high-school index for all the students from each particular school are given in columns 2 and 3. In column 4 the actual grade point averages earned by students during their first year of college are listed. The grade point averages in Column 5 were predicted from a regression equation. For the individual, the error of estimate is so high that knowledge of the student's high-school index does not enable one to predict with any degree of accuracy the degree of success he may attain in college. The differences in grades actually obtained and grades predicted by means of the regression equation are shown in the last column of the table. The range of these differences is all the way from +1.43 (one student from a small private school) to −0.84 (one student from a small public school). A study of this column of differences reveals that students from the largest high schools and a majority of the private schools on the whole make better grades in college than predicted for them on the basis of the regression equation.

TABLE II.—A COMPARISON OF THE ACTUAL AND PREDICTED UNIVERSITY GRADES BY TOWNS

Town	No. of Students	H-S Index	Grades in U.	Predicted Grades	Difference
1. Adjuntas	11	2.92	2.21	2.26	−0.05
2. Aguada	6	3.14	1.64	2.42	−0.78
3. Aguadilla	75	2.68	2.19	2.08	+0.11
4. Aibonito	20	2.81	1.86	2.16	−0.30
5. Arecibo	99	2.59	2.00	2.02	−0.02
6. Bayamón	128	2.67	1.95	2.05	−0.10
7. Cabo Rojo	39	2.50	1.96	1.98	−0.02
8. Caguas	149	2.72	2.17	2.13	+0.04
9. Canóvanas	18	2.61	1.41	1.96	−0.55
10. Carolina	71	2.71	1.81	2.09	−0.28
11. Cayey	52	2.65	2.25	2.04	+0.21
12. Ciales	11	2.76	2.33	2.13	+0.20
13. Coamo	41	2.93	1.87	2.28	−0.41
14. Comerío	16	2.87	1.92	2.18	−0.26
15. Corozal	13	3.09	1.96	2.36	−0.40
16. Fajardo	50	2.57	1.94	1.97	−0.03
17. Guánica	10	2.45	1.62	1.86	−0.24
18. Guayama	62	2.70	1.93	2.07	−0.14
19. Hormigueros	7	2.74	1.59	2.19	−0.60
20. Humacao	43	2.71	2.02	2.09	−0.07
21. Isabela	23	2.69	1.92	2.12	−0.20
22. Jayuya	10	2.54	1.82	1.99	−0.17
23. Juana Díaz	1	2.00	1.18	1.91	−0.73
24. Juncos	15	2.53	1.91	1.95	−0.04
25. Lares	36	2.65	1.97	2.02	−0.05
26. Manatí	59	2.82	1.96	2.17	−0.21
27. Mayagüez	197	2.42	2.04	1.99	+0.05
28. Orocovis	1	3.39	2.04	2.88	−0.84
29. Ponce	166	2.66	2.22	2.07	+0.15
30. Río Grande	24	2.83	1.58	2.17	−0.59
31. Río Piedras (U.P.R.)	104	2.70	2.39	1.99	+0.40
32. R. Piedras High School	151	2.60	1.97	2.01	−0.04
33. R. Piedras (Montalvo)	18	2.40	1.57	1.82	−0.25
34. Salinas	27	2.77	1.87	2.14	−0.27
35. San Germán	37	2.62	2.18	2.11	+0.07
36. S. Juan (Superior Central)	412	2.68	2.25	2.08	+0.17
37. S. Juan (Superior Nocturna)	9	2.72	2.39	2.04	+0.35
38. S. Juan (Negoc. Adultos)	4	2.81	2.37	2.06	+0.31
39. San Sebastián	33	2.76	1.86	2.12	−0.26
40. Toa Baja	16	2.43	1.77	1.95	−0.18
41. Utúado	48	2.70	1.87	2.10	−0.23
42. Vega Baja	37	2.77	1.95	2.22	−0.27

TABLE II.—(*Continued*)

Town	No. of Students	H-S Index	Grades in U.	Predicted Grades	Difference
43. Vieques	16	2.38	1.56	1.87	−0.31
44. Yabucoa	16	2.70	1.98	2.04	−0.06
45. Yauco	52	2.63	2.07	2.11	−0.04
Private Schools					
46. St. Joseph (Aibonito)	1	2.37	1.30	1.73	−0.43
47. San Felipe (Arecibo)	2	2.87	1.98	2.13	−0.15
48. Ac. Bautista (Barranquitas)	9	2.66	1.79	2.64	−0.85
49. Mons. Willinger (Barranquitas)	17	2.83	1.95	2.15	−0.20
50. Santa Rosa (Bayamón)	7	2.81	1.52	2.25	−0.73
51. Presbyterian (Lajas)	2	3.57	2.03	2.55	−0.52
52. San Luis (Lajas)	11	2.74	1.55	2.08	−0.53
53. Inmaculada (Manatí)	1	2.45	2.00	1.79	+0.21
54. Inmaculada (Mayagüez)	12	2.81	2.65	2.30	+0.35
55. Santa Teresa (Naranjito)	5	2.45	2.10	2.13	−0.03
56. Col. Ponceño de Varones	12	2.67	2.23	2.15	+0.08
57. Liceo Ponceño	4	2.84	2.71	2.21	+0.50
58. Modern Business College	2	2.04	1.83	1.48	+0.35
59. La Milagrosa (R. Piedras)	5	2.37	1.79	1.81	−0.02
60. San Antonio (R. Piedras)	9	2.53	1.69	1.89	−0.20
61. San Jorge (R. Piedras)	1	2.31	1.71	1.88	−0.17
62. San José (R. Piedras)	7	2.00	1.64	1.93	−0.29
63. P.R. High School of Commerce (Río Piedras)	1	2.03	2.90	1.47	+1.43
64. San José (San Germán)	5	3.47	2.51	2.50	+0.01
65. Acad. Católica (S. Juan)	5	3.06	2.18	2.32	−0.14
66. La Inmaculada (Santurce)	5	2.91	2.22	2.18	+0.04
67. Col. Puertorriqueño (Santurce)	12	2.77	2.56	2.08	+0.48
68. Perpetuo Socorro (Santurce)	10	3.31	2.36	2.57	−0.21
69. Sagrado Corazón (Santurce)	2	2.39	2.52	1.84	+0.68
70. Blanche Kellogg (Santurce)	3	2.57	2.34	1.94	+0.40
71. Santa Rita (S. Sebastián)	7	3.01	1.87	2.25	−0.38
72. Holy Rosary (Yabucoa)	2	2.73	2.12	2.19	−0.07
73. Holy Rosary (Yauco)	2	2.39	2.28	2.13	+0.15

Students from the smaller high schools do not attain the level expected of them.

This fact is more forcefully presented in Table III where nine of the larger and twenty-three of the smaller public high schools are listed and tabulated on the basis of the weighted differences between grades actually obtained in college and grades predicted. In column 1 intervals in terms of 'difference' units are set up. The number of students and the actual weighted 'differences' are

TABLE III.—DISTRIBUTION OF THE DIFFERENCES BETWEEN COLLEGE GRADES OBTAINED AND COLLEGE GRADES PREDICTED FOR NINE LARGE AND TWENTY-THREE SMALL HIGH SCHOOLS

Intervals	High School	N	Difference
0.20– 0.29	Ciales	11	0.20
0.10– 0.19	San Juan*	412	0.17
	Ponce*	166	0.15
0.00– 0.09	Mayagüez*	197	0.05
	Caguas*	149	0.04
−0.10−−0.01	Arecibo*	99	−0.02
	Cabo Rojo	39	−0.02
	Juncos	15	−0.04
	Río Piedras*	151	−0.04
	Yauco	52	−0.04
	Adjūntas	11	−0.05
	Lares	36	−0.05
	Yabucoa	16	−0.06
	Humacao*	43	−0.07
	Bayamón*	128	−0.10
−0.20−−0.11	Guayama*	62	−0.14
	Jayuya	10	−0.17
	Toa Baja	16	−0.18
	Isabela	23	−0.20
−0.30−−0.21	Utuado	48	−0.23
	Guánica	10	−0.24
	San Sebastián	33	−0.26
	Comerío	16	−0.26
	Salinas	27	−0.27
	Aibonito	20	−0.30
−0.40−−0.31	Vieques	16	−0.31
	Corozal	13	−0.40
−0.50−−0.41	Coamo	41	−0.41
−0.60−−0.51	Canóvanas	18	−0.55
	Río Grande	24	−0.59
	Hormigueros	7	−0.60
−0.70−−0.61			
−0.80−−0.71	Aguada	6	−0.78

* High schools with enrollments of 500 or more.

given in the last two columns. An asterisk has been placed by the name of each high school with a total enrollment of five hundred or over. The other twenty-three high schools included in the table have enrollments under three hundred.

It will be noticed that the large high schools are high in the distribution and that the small schools are low. This means that in general the students from the large high schools made indices in the colleges as good as, or better than, those predicted for them, but that the students from most of the small high schools actually earned college indices lower than those predicted for them on the basis of their high-school grades. Students with the same letter grades from small and large high schools do not have, it seems, the same degree of training and preparation for college work. High-school indices from various high schools and based on the system of letter grades in use do not appear to be equivalent.

CONCLUSIONS

On the basis of an analysis of the data used in this study, the following conclusions seem justified:

1) The high-school academic index is not very reliable as a criterion for selecting college students.

2) There is a positive relationship between high-school and college indices but the degree of correlation is so low that high-school indices cannot be used for predicting accurately college indices. The errors of estimate are high.

3) Students applying for admission to college who have the same high-school indices do not necessarily possess the same degree of training and preparation for college work.

4) An 'A' or 'B' grade in one high school is not necessarily equivalent to the corresponding 'A' or 'B' grade in another high school.

5) Students in general from the large high schools (enrollment of five hundred or more) make higher indices in college than do students from small high schools (enrollment of three hundred or less) with the same high-school indices.

6) The chances for success in college based on academic record are greater for students from the large high schools than for students from the small high schools with the same high-school indices.

7) Students with a 'C' average from large high schools make on the whole higher grades in college than do 'B' students from small schools.

8) The indications are that a grave injustice may be done students from large high schools with a high-school index slightly below that required for admission to college. They may be in many instances better prepared for college work than are students from small high schools who are admitted to college because they have high-school indices a few points higher.

RECOMMENDATIONS

The purposes of this study have been realized. The data have been collected, systematized, analyzed, and interpreted. Conclusions which seem to be well justified by the objective techniques employed have been listed. This paper should not end here, however. Certain suggestions of a constructive nature are in order. These recommendations concern the general problem of providing some kind of a screening process whereby those selected for college training shall be the individuals from among all the applicants for admission who may profit most from college work. Such a screening process would result in the admission of those students who have the greatest chance of succeeding at the college level. The recommendations offered here concerning this general problem may be conveniently classified under three main heads: (1) Those designed to improve the reliability of high school indices as a criterion for selection (1–5 inclusive); (2) those designed to utilize other factors as criteria for college entrance (6 and 7); and, finally, (3) those designed to extend and expand the educational opportunities and facilities of the Island above the high-school level (8–10 inclusive). Such extension and expansion would tend to change the nature of the selection process itself.

1) Improve instruction in all the public schools of the Island. This involves the provisions for more adequate school facilities and superior teachers.

2) Improve instruction, especially in the small high schools in order that graduates of these schools may be equally prepared (as graduates from large high schools) for college work.

3) A more uniform and standardized system of assigning school marks might be worked out.

4) Consolidate, in so far as is feasible, the small high schools of the Island. Distances within the Island are short and children may be readily transported by bus. Such consolidations would serve to place the advantages of large high schools within reach of all.

5) The Department of Education and the University of Puerto Rico should collaborate in establishing a division of Measurements and Guidance. Such a division would work throughout the year in the preparation of tests and other measuring instruments and in the development of techniques for student guidance. A comprehensive testing program would be set up for the public schools of the entire Island. The test results would be cumulative for each child and would be valuable information which might be used for a variety of purposes. Among other things the test results could be used as another criterion for admission to college and for the purpose of assisting students in choosing the courses in college from which they might derive the greatest benefit. If this were possible, the money which may be spent in giving students instruction in the lines of study for which they are not suited, could be used to greater advantage.

It seems reasonable to assume that the very high mortality rate which now exists might be considerably reduced by a good guidance program. From the economic point of view we should endeavor to hold students at the college level until they finish their studies. The per capita expenditure at the college level is very great. Moreover, mortality at this level involves individuals of high achievement and great capacity for learning. This in itself presents a problem of paramount importance since our country needs people with more training and ability in different fields of education to help in the solution of its social and economic problems.

6) High-school indices should not be used as a sole criterion for selection of college students.

7) Standardized objective tests of various kinds should be utilized as a screening device. Comprehensive tests designed to measure academic aptitude, reading proficiency (vocabulary and comprehension), achievement, and certain aspects of personality development would yield information of great value for all concerned. Such test results are essential for educational, vocational, and personal guidance. Many such tests are already

available and new ones could be constructed to meet the specific needs of Puerto Ricans.

8) Extend and expand the facilities of the University of Puerto Rico and all of its branches. The University should be prepared to meet the needs of a student body of from 15,000 to 20,000 within a few years. As a senior university and graduate school, it should be on a par with the best of them.

9) A system of public junior colleges offering various types of two-year terminal courses should be established. The Island might be divided into four or five regions with a public junior college located in each and accessible to the high-school graduates of that region.

10) Public business, technical, and vocational schools of various kinds to meet the needs of the people of the Island might be provided after a study is made to determine the needs of the people along these lines.

In conclusion, it should be stated with considerable emphasis, that additional research on this and related problems is sorely needed.

UNIVERSITY OF PUERTO RICO LIBRARY